UTILITY APPLICATION

UNDER 37 CFR § 1.53(B) (2)

TITLE: DNA SEQUENCES FROM *STAPHYLOCOCCUS AUREUS* BACTERRIOPHAGE 44AHJD THAT ENCODE ANTI-MICROBIAL POLYPEPTIDES

APPLICANT: Jerry Pelletier, Philippe Gros, Micheal Dubow

Correspondence Enclosed:

Utility Application (2 pgs); Cover Sheet (1 pg); Specification (88 pgs); Claims (9 pgs); Abstract (1 pg); Table 1- 8 (58 pgs); Figures 1-10 (16 pgs); and Return Postcard

"EXPRESS MAIL" Mailing Label Number EL675944637US Date of Deposit November 30, 2000 I hereby certify under 37 CFR §1.10 that this correspondence is being deposited with the United States Postal Service as "Express Mail Post Office to Addressee" with sufficient postage on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Nancy Overly-Walker

DNA SEQUENCES FROM *STAPHYLOCOCCUS AUREUS* BACTERIOPHAGE 44AHJD THAT ENCODE ANTI-MICROBIAL POLYPEPTIDES

RELATED APPLICATIONS

This application claims the benefit of United States Provisional Application Pelletier, *et al.*, 60/168,777 filed December 1, 1999, which is hereby incorporated by reference in its entirety, including drawings.

BACKGROUND OF THE INVENTION

This invention relates to the identification of antimicrobial agents and of microbial targets of such agents, and in particular to the isolation of bacteriophage DNA sequences, and their translated protein products, showing anti-microbial activity. The DNA sequences can be expressed in expression vectors. These expression constructs and the proteins produced therefrom can be used for a variety of purposes including therapeutic methods and identification of microbial targets.

The following description is provided to assist the understanding of the reader. None of the information provided or references cited is admitted to be prior art to the present invention.

The frequency and spectrum of antibiotic-resistant infections have, in recent years, increased in both the hospital and community. Certain infections have become essentially untreatable and are growing to epidemic proportions in the developing world as well as in institutional settings in the developed world. The staggering spread of antibiotic resistance in pathogenic bacteria has been attributed to microbial genetic characteristics, widespread use of antibiotic drugs and changes in society that enhance the transmission of drug-resistant organisms (for a review, see Cohen, 1992). This spread of drug resistant microbes is leading to ever-increasing morbidity, mortality and health-care costs.

There are over 160 antibiotics currently available for treatment of microbial infections, all based on a few basic chemical structures and targeting a small number of metabolic pathways: bacterial cell wall synthesis, protein synthesis, and DNA replication.

Despite all these antibiotics, a person could succumb to an infection as a result of a resistant bacterial infection. Resistance now reaches all classes of antibiotics currently in use, including: β-lactams, fluoroquinolones, aminoglycosides, macrolide peptides, chloramphenicol, tetracyclines, rifampicin, folate inhibitors, glycopeptides, and mupirocin. There is thus a need for new antibiotics, and this need will not subside given the ability bacteria have to overcome each new agent synthesized. It is also likely that targeting new pathways will play an important role in discovery of these new antibiotics. In fact, a number of crucial cellular pathways, such as secretion, cell division, and many metabolic functions, remain untargeted today.

Most major pharmaceutical companies have on-going drug discovery programs for novel anti-microbials. These are based on screens for small molecule inhibitors (e.g., natural products, bacterial culture media, libraries of small molecules, combinatorial chemistry) of crucial metabolic pathways of the micro-organism of interest. The screening process is largely for cytotoxic compounds and in most cases is not based on a known mechanism of action of the compounds. Classical drug screening programs are being exhausted and many of these pharmaceutical companies are looking towards rational drug design programs. Several small to mid-size biotechnology companies, as well as large pharmaceutical companies, have developed systematic high-throughput sequencing programs to decipher the genetic code of specific micro-organisms of interest. The goal is to identify, through sequencing, unique biochemical pathways or intermediates that are unique to the microorganism. Knowledge of the function of these bacterial genes, may form the rationale for a drug discovery program based on the mechanism of action of the identified enzymes/proteins. However, one of the most critical steps in this approach is the ascertainment that the identified proteins and biochemical pathways are 1) non-redundant and essential for bacterial survival, and 2) constitute suitable and accessible targets for drug discovery. These two issues are not easily addressed since to date, 18 prokaryotic genomes have been sequenced and 200 sequenced genomes are expected by the year 2000. For a majority of the sequenced genomes, less than 50% of the open reading frames (ORFs) have been linked to a known function. Even with the genome of Escherichia coli (*E. coli*), the most extensively studied bacterium, less than two-thirds of the annotated protein coding

genes showed significant similarity to genes with ascribed functions (Rusterholtz and Pohlschroder, 1999). Thus considerable work must be undertaken to identify appropriate bacterial targets for drug screening.

SUMMARY OF THE INVENTION

The present invention is based on the identification of, and demonstration that, specific DNA sequences of a bacteriophage, when introduced into a host bacterium can kill, or inhibit growth, of the host. Thus, these DNA sequences are anti-microbial agents. Information based on these DNA sequences can be utilized to develop peptide mimetics that can also function also as anti-microbials. The identification of the host bacterial proteins, targeted by the anti-microbial bacteriophage DNA sequences, can provide novel targets for drug design, compound screening or determination of new domains on an already known target.

In this regard, the terms "inhibit", "inhibition", "inhibitory", and "inhibitor" all refer to a function of reducing a biological activity or function. Such reduction in activity or function can, for example, be in connection with a cellular component (e.g., an enzyme), or in connection with a cellular process (e.g., synthesis of a particular protein), or in connection with an overall process of a cell (e.g., cell growth). In reference to cell growth, the inhibitory effects may be bactericidal (killing of bacterial cells) or bacteriostatic (i.e., stopping or at least slowing bacterial cell growth). The latter slows or prevents cell growth such that fewer cells of the strain are produced relative to uninhibited cells over a given time period. From a molecular standpoint, such inhibition may equate with a reduction in the level of, or elimination of, the transcription and/or translation of a specific bacterial target(s), or reduction or elimination of activity of a particular target biomolecule.

In a first aspect the invention provides methods for identifying a target for antibacterial agents by identifying the bacterial target(s) of at least one inhibitory gene product, e.g., protein from ORFs 12 and 25, of bacteriophage 44AHJD or a homologous product. Such identification allows the development of antibacterial agents active on such targets. Preferred embodiments for identifying such targets involve the identification of binding of target and phage ORF products to one another. The target molecule may be a

bacterial protein or other bacterial biomolecule, e.g., a nucleoprotein, a nucleic acid, a lipid or lipid-containing molecule, a nucleoside or nucleoside derivative, a polysaccharide or polysaccharide-containing molecule, or a peptidoglycan. The phage ORF products may be subportions of a larger ORF product that also binds the host target. Exemplary approaches are described below in the Detailed Description.

Additionally, the invention provides methods for identifying targets for antibacterial agents by identifying homologs of a *Staphylococcus aureus* target of a bacteriophage 44 AHJD ORFs 12 or 25 product. Such homologs may be utilized in the various aspects and embodiments described herein.

The term "fragment" refers to a portion and/or a segment of a larger molecule or assembly. For proteins, the term "fragment" refers to a molecule which includes at least 5 contiguous amino acids from the reference polypeptide or protein, preferably at least 6, 8, 10, 12, 15, 20, 30, 50 or more contiguous amino acids. In connection with oligo- or polynucleotides, the term "fragment" refers to a molecule which includes at least 15 contiguous nucleotides from a reference polynucleotide, preferably at least 18, 21, 24, 30, 36, 45, 60, 90, 150, or more contiguous nucleotides. Also in preferred embodiments, the fragment has a length in a range with the minimum as described above and a maximum which is no more than 90 % of the length (or contains that percent of the contiguous amino acids or nucleotides) of the larger molecule (e.g., of the specified ORF), in other embodiments, the upper limit is no more than 60, 70, or 80% of the length of the larger molecule.

Stating that an agent, compound or test compound is "active on" a particular cellular target, such as the product of a particular gene, means that the target is an important part of a cellular pathway which includes that target and that the agent acts on that pathway. Thus, in some cases the agent may act on a component upstream or downstream of the stated target, including a regulator of that pathway or a component of that pathway. In general, an antibacterial agent is active on an essential cellular function, often on a product of an essential gene.

By "essential", in connection with a gene or gene product, is meant that the host cannot survive without, or is significantly growth compromised, in the absence or depletion

of functional product. An "essential gene" is thus one that encodes a product that is beneficial, or preferably necessary, for cellular growth in vitro in a medium appropriate for growth of a strain having a wild-type allele corresponding to the particular gene in question. Therefore, if an essential gene is inactivated or inhibited, that cell will grow significantly more slowly or even not at all. Preferably growth of a strain in which such a gene has been inactivated will be less than 20%, more preferably less than 10%, most preferably less than 5% of the growth rate of the wild-type, or not at all, in the growth medium. Preferably, in the absence of activity provided by a product of the gene, the cell will not grow at all or will be non-viable, at least under culture conditions similar to normal *in vivo* growth conditions. For example, absence of the biological activity of certain enzymes involved in bacterial cell wall synthesis can result in the lysis of cells under normal osmotic conditions, even though protoplasts can be maintained under controlled osmotic conditions. Preferably, but not necessarily, if such a gene is inhibited, e.g., with an antibacterial agent or a phage product, the growth rate of the inhibited bacteria will be less than 50%, more preferably less than 30%, still more preferably less than 20%, and most preferably less than 10% of the growth rate of the uninhibited bacteria. As recognized by those skilled in the art, the degree of growth inhibition will generally depend on the concentration of the inhibitory agent. In the context of the invention, essential genes are generally the preferred targets of antimicrobial agents. Essential genes can encode target molecules directly or can encode a product involved in the production, modification, or maintenance of a target molecule.

A "target" refers to a biomolecule that can be acted on by an exogenous agent, thereby modulating, preferably inhibiting, growth or viability of a cell. In most cases such a target will be a nucleic acid sequence or molecule, or a polypeptide or protein. However, other types of biomolecules can also be targets, e.g., membrane lipids and cell wall structural components.

The term "bacterium" refers to a single bacterial strain, and includes a single cell, and a plurality or population of cells of that strain unless clearly indicated to the contrary. In reference to bacteria or bacteriophage, the term "strain" refers to bacteria or phage having a particular genetic content. The genetic content includes genomic content as well as recombinant vectors. Thus, for example, two otherwise identical bacterial cells would

represent different strains if each contained a vector, e.g., a plasmid, with different phage ORF inserts.

In the context of the phage nucleic acid sequences, e.g., gene sequences, of this invention, the terms "homolog" and "homologous" denote nucleotide sequences from different bacteria or phage strains or species or from other types of organisms that have significantly related nucleotide sequences, and consequently significantly related encoded gene products, preferably having related function. Homologous gene sequences or coding sequences have at least 70% sequence identity (as defined by the maximal base match in a computer-generated alignment of two or more nucleic acid sequences) over at least one sequence window of 48 nucleotides (or at least 99, 150, 200, or even the entire ORF or other sequence of interest), more preferably at least 80 or 85%, still more preferably at least 90%, and most preferably at least 95%. The polypeptide products of homologous genes have at least 35% amino acid sequence identity over at least one sequence window of 18 amino acid residues (or 24, 30, 33, 50, 100, or an entire polypeptide), more preferably at least 40%, still more preferably at least 50% or 60%, and most preferably at least 70%, 80%, or 90%. Preferably, the homologous gene product is also a functional homolog, meaning that the homolog will functionally complement one or more biological activities of the product being compared. For nucleotide or amino acid sequence comparisons where a homology is defined by a % sequence identity, the percentage is determined using BLAST programs (with default parameters (Altschul et al., 1997, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acid Res. 25:3389-3402). Any of a variety of algorithms known in the art which provide comparable results can also be used, preferably using default parameters. Performance characteristics for three different algorithms in homology searching is described in Salamov et al., 1999, "Combining sensitive database searches with multiple intermediates to detect distant homologues." *Protein Eng.* 12:95-100. Another exemplary program package is the GCG™ package from the University of Wisconsin.

Homologs may also or in addition be characterized by the ability of two complementary nucleic acid strands to hybridize to each other under appropriately stringent conditions. Hybridizations are typically and preferably conducted with probe-length nucleic

acid molecules, preferably 20-100 nucleotides in length. Those skilled in the art understand how to estimate and adjust the stringency of hybridization conditions such that sequences having at least a desired level of complementarity will stably hybridize, while those having lower complementarity will not. For examples of hybridization conditions and parameters, see, e.g.,. Maniatis, T. *et al.* (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor University Press, Cold Spring, N.Y.; Ausubel, F.M. *et al.* (1994) Current Protocols in Molecular Biology. John Wiley & Sons, Secaucus, N.J. Homologs and homologous gene sequences may thus be identified using any nucleic acid sequence of interest, including the phage ORFs and bacterial target genes of the present invention.

A typical hybridization, for example, utilizes, besides the labeled probe of interest, a salt solution such as 6X SSC (NaCl and Sodium Citrate base) to stabilize nucleic acid strand interaction, a mild detergent such as 0.5% SDS, together with other typical additives such as Denhardt's solution and salmon sperm DNA. The solution is added to the immobilized sequence to be probed and incubated at suitable temperatures to preferably permit specific binding while minimizing nonspecific binding. The temperature of the incubations and ensuing washes is critical to the success and clarity of the hybridization. Stringent conditions employ relatively higher temperatures, lower salt concentrations, and/or more detergent than do non-stringent conditions. Hybridization temperatures also depend on the length, complementarity level, and nature (i.e., "GC content") of the sequences to be tested. Typical stringent hybridizations and washes are conducted at temperatures of at least 40°C, while lower stringency hybridizations and washes are typically conducted at 37°C down to room temperature (~25°C). One of ordinary skill in the art is aware that these conditions may vary according to the parameters indicated above, and that certain additives such as formamide and dextran sulphate may also be added to affect the conditions.

By "stringent hybridization conditions" is meant hybridization conditions at least as stringent as the following: hybridization in 50% formamide, 5X SSC, 50 mM NaH2PO4, pH 6.8, 0.5% SDS, 0.1 mg/mL sonicated salmon sperm DNA, and 5X Denhart's solution at 42°C overnight; washing with 2X SSC, 0.1% SDS at 45°C; and washing with 0.2X SSC, 0.1% SDS at 45°C. In another example, stringent hybridization conditions should not allow

for hybridization of two nucleic acids which differ over a stretch of 20 contiguous nucleotides by more than two bases.

Homologous nucleotide sequences will distinguishably hybridize with a reference sequence with up to three mismatches in ten (i.e., at least 70% base match in two sequences of equal length). Preferably, the allowable mismatch level is up to two mismatches in 10, or up to one mismatch in ten, more preferably up to one mismatch in twenty. (Those ratios can, of course, be applied to larger sequences.)

Preferred embodiments involve identification of binding between ORF product and bacterial cellular component that include methods for distinguishing bound molecules, for example, affinity chromatography, immunoprecipitation, crosslinking, and/or genetic screen methods that permit protein:protein interactions to be monitored. One of skill in the art is familiar with these techniques and common materials utilized (see, e.g., Coligan, J. *et al.* (eds.) (1995) Current Protocols in Protein Science, John Wiley & Sons, Secaucus, N.J.).

Genetic screening for the identification of protein:protein interactions typically involves the co-introduction of both a chimeric bait nucleic acid sequence (here, the phage ORF to be tested) and a chimeric target nucleic acid sequence that, when co-expressed and having affinity for one another in a host cell, stimulate reporter gene expression to indicate the relationship. A "positive" can thus suggest a potential inhibitory effect in bacteria. This is discussed in further detail in the Detailed Description section below. In this way, new bacterial targets can be identified that are inhibited by specific phage ORF products or derivatives, fragments, mimetics, or other molecules.

Other embodiments involve the identification and/or utilization of a target which is mutated at the site of phage 44AHJD protein interaction but still functional in the cell by virtue of their host's relatively unresponsive nature in the presence of expression of ORFs previously identified as inhibitory to the non-mutant or wild-type strain. Such mutants have the effect of protecting the host from an inhibition that would otherwise occur (e.g., by competing for binding with the phage ORF product) and indirectly allow identification of the precise responsible target. The identified target can then be used, e.g., for follow-up studies and anti-microbial development. In certain embodiments, rescue from inhibition occurs under conditions in which a bacterial target or mutant target is highly expressed. This

is performed, for example, through coupling of the sequence with regulatory element promoters, e.g., as known in the art, which regulate expression at levels higher than wild-type, e.g., at a level sufficiently higher that the inhibitor can be competitively bound to the highly expressed target such that the bacterium is detectably less inhibited.

Identification of the bacterial target can involve identification of a phage-specific site of action. This can involve a newly identified target, or a target where the phage site of action differs from the site of action of a previously known antibacterial agent or inhibitor. For example, phage T7 genes 0.7 and 2.0 target the host RNA polymerase, which is also the cellular target for the antibacterial agent, rifampin. To the extent that a phage product is found to act at a different site than previously described inhibitors, aspects of the present invention can utilize those new, phage-specific sites for identification and use of new agents. The site of action can be identified by techniques known to those skilled in the art, for example, by mutational analysis, binding competition analysis, and/or other appropriate techniques.

Once a bacterial host target or mutant target sequence has been identified, it too can be conveniently sequenced, sequence analyzed (e.g., by computer), and the underlying gene(s), and corresponding translated product(s) further characterized. Preferred embodiments include such analysis and identification. Preferably such a target has not previously been identified as an appropriate target for antibacterial action.

Also in preferred embodiments in which the bacterial target is a polypeptide or nucleic acid molecule, the identification of a bacterial target of a phage ORF product or fragment includes identification of a cellular and/or biochemical function of the bacterial target. As understood by those skilled in the art, this can, for example, include identification of function by identification of homologous polypeptides or nucleic acid molecules having known function, or identification of the presence of known motifs or sequences corresponding to known function. Such identifications can be readily performed using sequence comparison computer software, such as the BLAST programs and similar other programs and sequence and motif databases.

In embodiments involving expression of a phage ORF in a bacterial strain, in preferred embodiments that expression is inducible. By "inducible" is meant that expression

is absent or occurs at a low level until the occurrence of an appropriate environmental stimulus provides otherwise. For the present invention such induction is preferably controlled by an artificial environmental change, such as by contacting a bacterial strain population with an inducing compound (i.e., an inducer). However, induction could also occur, for example, in response to build-up of a compound produced by the bacteria in the bacterial culture, e.g., in the medium. As uncontrolled or constitutive expression of inhibitory ORFs can severely compromise bacteria to the point of eradication, such expression is therefore undesirable in many cases because it would prevent effective evaluation of the strain and inhibitor being studied. For example, such uncontrolled expression could prevent any growth of the strain following insertion of a recombinant ORF, thus preventing determination of effective transfection or transformation. A controlled or inducible expression is therefore advantageous and is generally provided through the provision of suitable regulatory elements, e.g., promoter/operator sequences that can be conveniently transcriptionally linked to a coding sequence to be evaluated. In most cases, the vector will also contain sequences suitable for efficient replication of the vector in the same or different host cells and/or sequences allowing selection of cells containing the vector, i.e., "selectable markers." Further, preferred vectors include convenient primer sequences flanking the cloning region from which PCR and/or sequencing may be performed. In preferred embodiments where the purification of phage product is desired, preferably the bacterium or other cell type does not produce a target for the inhibitory product, or is otherwise resistant to the inhibitory product.

In preferred embodiments, the target of the phage ORF product or fragment is identified from a bacterial animal pathogen, preferably a mammalian pathogen, more preferably a human pathogen, and is preferably a gene or gene product of such a pathogen. Also in preferred embodiments, the target is a gene or gene product, where the sequence of the target is homologous to a gene or gene product from such a pathogen as identified above.

As used herein, the term "mammal" has its usual biological meaning, and particularly includes bovines, swine, dogs, cats, and humans.

Other aspects of the invention provide isolated, purified, or enriched specific phage nucleic acid and amino acid sequences, subsequences, and homologs thereof preferably from

or corresponding to ORFs 12 and 25, from bacteriophage 44AHJD (*Staphylococcus aureus* host bacterium). Such nucleotide sequences are at least 15 nucleotides in length, preferably at least 18, 21, 24, or 27 nucleotides in length, more preferably at least 30, 50, or 90 nucleotides in length. In certain embodiments, longer nucleic acids are preferred, for example those of at least 120, 150, 200, 300, 600, 900 or more nucleotides. Such sequences can, for example, be amplification oligonucleotides (e.g., PCR primers), oligonucleotide probes, sequences encoding a portion or all of a phage-encoded protein, or a fragment or all of a phage-encoded protein. In preferred embodiments, the nucleic acid sequence or amino acid sequence contains a sequence which has a lower length as specified above, and an upper-length limit which is no more than 50, 60, 70, 80, or 90% of the length of the full-length ORF or ORF product. The upper-length limit can also be expressed in terms of the number of base pairs of the ORF (coding region).

As it is recognized that alternate codons will encode the same amino acid for most amino acids due to the degeneracy of the genetic code, the sequences of this aspect includes nucleic acid sequences utilizing such alternate codon usage for one or more codons of a coding sequence. For example, all four nucleic acid sequences GCT, GCC, GCA, and GCG encode the amino acid, alanine. Therefore, if for an amino acid there exists an average of three codons, a polypeptide of 100 amino acids in length will, on average, be encoded by 3^{100}, or 5×10^{47}, nucleic acid sequences. Thus, a nucleic acid sequence can be modified (e.g., a nucleic acid sequence from a phage as specified above) to form a second nucleic acid sequence encoding the same polypeptide as encoded by the first nucleic acid sequence using routine procedures and without undue experimentation. Thus, all possible nucleic acid sequences that encode the amino acid sequences encoded by the phage 44AHJD ORFs 12 and 25, as if all were written out in full, taking into account the codon usage, especially that preferred in the host bacterium.

The alternate codon descriptions are available in common textbooks, for example, Stryer, BIOCHEMISTRY 3rd ed., and Lehninger, BIOCHEMISTRY 3rd ed. Codon preference tables for various types of organisms are available in the literature. Because of the number of sequence variations involving alternate codon usage, for the sake of brevity, individual sequences are not separately listed herein. Instead the alternate sequences are

described by reference to the natural sequence with replacement of one or more (up to all) of the degenerate codons with alternate codons from the alternate codon table (Table 2), preferably with selection according to preferred codon usage for the normal host organism or a host organism in which a sequence is intended to be expressed. Those skilled in the art also understand how to alter the alternate codons to be used for expression in organisms where certain codons code differently than shown in the "universal" codon table.

For amino acid sequences, sequences contain at least 5 peptide-linked amino acid residues, and preferably at least 6, 7, 10, 15, 20, 30, or 40, amino acids having identical amino acid sequence as the same number of contiguous amino acid residues in a phage ORF 12 or 25 product. In some cases longer sequences may be preferred, for example, those of at least 50, 70, or 100 amino acids in length. In preferred embodiments, the sequence has bacteria-inhibiting function when expressed or otherwise present in a bacterial cell which is a host for the bacteriophage from which the sequence was derived.

By "isolated" in reference to a nucleic acid is meant that a naturally occurring sequence has been removed from its normal cellular (e.g., chromosomal) environment or is synthesized in a non-natural environment (e.g., artificially synthesized). Thus, the sequence may be in a cell-free solution or placed in a different cellular environment. The term does not imply that the sequence is the only nucleotide chain present, but that it is essentially free (about 90-95% pure at least) of non-nucleotide material naturally associated with it, and thus is distinguished from isolated chromosomes.

The term "enriched" means that the specific DNA or RNA sequence constitutes a significantly higher fraction (2-5 fold) of the total DNA or RNA present in the cells or solution of interest than in normal or diseased cells or in cells from which the sequence was originally taken. This could be caused by a person by preferential reduction in the amount of other DNA or RNA present, or by a preferential increase in the amount of the specific DNA or RNA sequence, or by a combination of the two. However, it should be noted that enriched does not imply that there are no other DNA or RNA sequences present, just that the relative amount of the sequence of interest has been significantly increased.

The term "significant" is used to indicate that the level of increase is useful to the person making such an increase and an increase relative to other nucleic acids of about at

least 2-fold, more preferably at least 5- to 10-fold or even more. The term also does not imply that there is no DNA or RNA from other sources. The other source DNA may, for example, comprise DNA from a yeast or bacterial genome, or a cloning vector such as pUC19. This term distinguishes from naturally occurring events, such as viral infection, or tumor type growths, in which the level of one mRNA may be naturally increased relative to other species of mRNA. That is, the term is meant to cover only those situations in which a person has intervened to elevate the proportion of the desired nucleic acid.

It is also advantageous for some purposes that a nucleotide sequence be in purified form. The term "purified" in reference to nucleic acid does not require absolute purity (such as a homogeneous preparation). Instead, it represents an indication that the sequence is relatively more pure than in the natural environment (compared to the natural level, this level should be at least 2-5 fold greater, e.g., in terms of mg/mL). Individual clones isolated from a genomic or cDNA library may be purified to electrophoretic homogeneity. The claimed DNA molecules obtained from these clones could be obtained directly from total DNA or from total RNA. cDNA clones are not naturally occurring, but rather are preferably obtained via manipulation of a partially purified naturally occurring substance (messenger RNA). The construction of a cDNA library from mRNA involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection of the cells carrying the cDNA library. Thus, the process which includes the construction of a cDNA library from mRNA and isolation of distinct cDNA clones yields an approximately 106-fold purification of the native message. Thus, purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. A genomic library can be used in the same way and yields the same approximate levels of purification.

The terms "isolated", "enriched", and "purified" with respect to the nucleic acids, above, may similarly be used to denote the relative purity and abundance of polypeptides (multimers of amino acids joined one to another by α-carboxyl:α-amino group (peptide) bonds). These, too, may be stored in, grown in, screened in, and selected from libraries using biochemical techniques familiar in the art. Such polypeptides may be natural, synthetic or chimeric and may be extracted using any of a variety of methods, such as antibody

immunoprecipitation, other "tagging" techniques, conventional chromatography and/or electrophoretic methods. Some of the above utilize the corresponding nucleic acid sequence.

As indicated above, aspects and embodiments of the invention are not limited to entire genes and proteins. The invention also provides and utilizes fragments and portions thereof, preferably those which are "active" in the inhibitory sense described above. Such peptides or oligopeptides and oligo or polynucleotides have preferred lengths as specified above for nucleic acid and amino acid sequences from phage; corresponding recombinant constructs can be made to express the encoded same. Also included are homologous sequences and fragments thereof.

The nucleotide and amino acid sequences identified herein are believed to be correct, however, certain sequences may contain a small percentage of errors, e.g., 1-5%. In the event that any of the sequences have errors, the corrected sequences can be readily provided by one skilled in the art using routine methods. For example, the nucleotide sequences can be confirmed or corrected by obtaining and culturing the relevant phage, and purifying phage genomic nucleic acids. A region or regions of interest can be amplified, e.g., by PCR from the appropriate genomic template, using primers based on the described sequence. The amplified regions can then be sequenced using any of the available methods (e.g., a dideoxy termination method, for example, using commercially available products). This can be done redundantly to provide the corrected sequence or to confirm that the described sequence is correct. Alternatively, a particular sequence or sequences can be identified and isolated as an insert or inserts in a phage genomic library and isolated, amplified, and sequenced by standard methods. Confirmation or correction of a nucleotide sequence for a phage gene provides an amino acid sequence of the encoded product by merely reading off the amino acid sequence according to the normal codon relationships and/or expressed in a standard expression system and the polypeptide product sequenced by standard techniques. The sequences described herein thus provide unique identification of the corresponding genes and other sequences, allowing those sequences to be used in the various aspects of the present invention. Confirmation of a phage ORF encoded amino acid sequence can also be confirmed by constructing a recombinant vector from which the ORF can be expressed in an

appropriate host (e.g., *E. coli*), purified, and sequenced by conventional protein sequencing methods.

In other aspects the invention provides recombinant vectors and cells harboring, one or more phage 44AHJD ORFs, preferably ORF 12 or 25 or portions thereof, or bacterial target sequences described herein, preferably where the phage or bacterial sequence is inserted in a recombinant vector. As understood by those skilled in the art, vectors may assume different forms, including, for example, plasmids, cosmids, and virus-based vectors. See, e.g., Maniatis, T. *et al.* (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor University Press, Cold Spring, N.Y.; See also, Ausubel, F.M. *et al.* (eds.) (1994) Current Protocols in Molecular Biology. John Wiley & Sons, Secaucus, N.J.

In preferred embodiments, the vectors will be expression vectors, preferably shuttle vectors that permit cloning, replication, and expression within bacteria. An "expression vector" is one having regulatory nucleotide sequences containing transcriptional and translational regulatory information that controls expression of the nucleotide sequence in a host cell. Preferably the vector is constructed to allow amplification from vector sequences flanking an insert locus. In certain embodiments, the expression vectors may additionally or alternatively support expression, and/or replication in animal, plant and/or yeast cells due to the presence of suitable regulatory sequences, e.g., promoters, enhancers, 3' stabilizing sequences, primer sequences, etc. In preferred embodiments, the promoters are inducible and specific for the system in which expression is desired, e.g., bacteria, animal, plant, or yeast. The vectors may optionally encode a "tag" sequence or sequences to facilitate protein purification or protein detection. Convenient restriction enzyme cloning sites and suitable selective marker(s) are also optionally included. Such selective markers can be, for example, antibiotic resistance markers or markers which supply an essential nutritive growth factor to an otherwise deficient mutant host, e.g., tryptophan, histidine, or leucine in the Yeast Two-Hybrid systems described below.

The term "recombinant vector" relates to a single- or double-stranded circular nucleic acid molecule that can be transfected into cells and replicated within or independently of a cell genome. A circular double-stranded nucleic acid molecule can be cut and thereby linearized upon treatment with appropriate restriction enzymes. An

assortment of nucleic acid vectors, restriction enzymes, and the knowledge of the nucleotide sequences cut by restriction enzymes are readily available to those skilled in the art. A nucleic acid molecule encoding a desired product can be inserted into a vector by cutting the vector with restriction enzymes and ligating the two pieces together. Preferably the vector is an expression vector, e.g., a shuttle expression vector as described above.

By " recombinant cell" is meant a cell possessing introduced or engineered nucleic acid sequences, e.g., as described above. The sequence may be in the form of or part of a vector or may be integrated into the host cell genome. Preferably the cell is a bacterial cell.

In preferred embodiments, the inserted nucleic acid sequence corresponding to at least a portion of a bacteriophage 44AHJD ORF, for example ORFs 12 and 25, gene product has a length as specified for the isolated purified or enriched nucleic acid sequences in an aspect above.

In another aspect, the invention also provides methods for identifying and/or screening compounds "active on" at least one bacterial target of a bacteriophage inhibitor protein or RNA, for example, corresponding to ORFs 12 or 25. Preferred embodiments involve contacting bacterial target proteins with a test compound, and determining whether the compound binds to or reduces the level of activity of the bacterial target, e.g., a bacterial protein. Preferably this is done in vivo under approximately physiological conditions. The compounds that can be used may be large or small, synthetic or natural, organic or inorganic, proteinaceous or non-proteinaceous. In preferred embodiments, the compound is a peptidomimetic, as described herein, a bacteriophage inhibitor protein or fragment or derivative thereof, preferably an "active portion", or a small molecule. In particular embodiments, the methods include the identification of bacterial targets as described above or otherwise described herein. Preferably the fragment of a bacteriophage inhibitor protein includes less than 80% of an intact bacteriophage inhibitor protein. Preferably, the at least one target includes a plurality of different targets of bacteriophage inhibitor proteins, preferably a plurality of different targets. The plurality of targets can be in or from a plurality of different bacteria, but preferably is from a single bacterial species.

In embodiments involving binding assays, preferably binding is to a fragment or portion of a bacterial target protein, where the fragment includes less than 90%, 80%, 70%,

60%, 50%, 40%, or 30% of an intact bacterial target protein. Preferably, the at least one bacterial target includes a plurality of different targets of bacteriophage inhibitor proteins, preferably a plurality of different targets. The plurality of targets can be in or from a plurality of different bacteria, but preferably is from a single bacterial species.

A "method of screening" refers to a method for evaluating a relevant activity or property of a large plurality of compounds, rather than just one or a few compounds. For example, a method of screening can be used to conveniently test at least 100, more preferably at least 1000, still more preferably at least 10,000, and most preferably at least 100,000 different compounds, or even more.

In the context of this invention, the term "small molecule" refers to compounds having molecular mass of less than 3000 Daltons, preferably less than 2000 or 1500, still more preferably less than 1000, and most preferably less than 600 Daltons. Preferably but not necessarily, a small molecule is not an oligopeptide.

In a related aspect or in preferred embodiments, the invention provides a method of screening for potential antibacterial agents by determining whether any of a plurality of compounds, preferably a plurality of small molecules, is active on at least one target of a bacteriophage inhibitor protein or RNA, for example, a target of ORF 12 or 25 gene product. Preferred embodiments include those described for the above aspect, including embodiments which involve determining whether one or more test compounds bind to or reduce the level of activity of a bacterial target, and embodiments which utilize a plurality of different targets as described above.

The identification of bacteria-inhibiting phage ORFs and their encoded products also provides a method for identifying an active portion of such an encoded product. This also provides a method for identifying a potential antibacterial agent by identifying such an active portion of a phage ORF or ORF product. In preferred embodiments, the identification of an active portion involves one or more of mutational analysis, deletion analysis, or analysis of fragments of such products. The method can also include determination of a 3-dimensional structure of an active portion, such as by analysis of crystal diffraction patterns. In further embodiments, the method involves constructing or synthesizing a peptidomimetic

compound, where the structure of the peptidomimetic compound corresponds to the structure of the active portion.

In this context, "corresponds" means that the peptidomimetic compound structure has sufficient similarities to the structure of the active portion that the peptidomimetic will interact with the same molecule as the phage protein and preferably will elicit at least one cellular response in common which relates to the inhibition of the cell by the phage protein.

The methods for identifying or screening for compounds or agents active on a bacterial target of a phage-encoded inhibitor can also involve identification of a phage-specific site of action on the target.

An "active portion" as used herein denotes an epitope, a catalytic or regulatory domain, or a fragment of a bacteriophage inhibitor protein that is responsible for, or a significant factor in, bacterial target inhibition. The active portion preferably may be removed from its contiguous sequences and, in isolation, still effect inhibition.

By "mimetic" is meant a compound structurally and functionally related to a reference compound that can be natural, synthetic, or chimeric. In terms of the present invention, a "peptidomimetic," for example, is a compound that mimics the activity-related aspects of the 3-dimensional structure of a peptide or polypeptide in a non-peptide compound, for example mimics the structure of a peptide or active portion of a phage- or bacterial ORF-encoded polypeptide.

A related aspect provides a method for inhibiting a bacterial cell by contacting the bacterial cell with a compound active on a bacterial target of a bacteriophage 44 AHJD inhibitor protein or RNA, preferably encoded by or corresponding to bacteriophage 44 AHJD ORF 12 or 25, where the target was uncharacterized. In preferred embodiments, the compound is such a protein, or a fragment or derivative thereof; a structural mimetic, e.g., a peptidomimetic, of such a protein or fragment; a small molecule; the contacting is performed in vitro, the contacting is performed in vivo in an infected or at risk organism, e.g., an animal such as a mammal or bird, for example, a human, or other mammal described herein, or in a plant.

In the context of this invention, the term "bacteriophage inhibitor protein" refers to a protein encoded by a bacteriophage nucleic acid sequence which inhibits bacterial function in a host bacterium. Thus, it is a bacteria-inhibiting phage product.

In the context of this invention, the phrase "contacting the bacterial cell with a compound active on a bacterial target of a bacteriophage inhibitor protein" or equivalent phrases refer to contacting with an isolated, purified, or enriched compound or a composition including such a compound, but specifically does not rely on contacting the bacterial cell with an intact naturally occurring phage which encodes the compound. Preferably no intact phage are involved in the contacting.

Related aspects provide methods for prophylactic or therapeutic treatment of a bacterial infection by administering to an infected, challenged or at risk organism a therapeutically or prophylactically effective amount of a compound active on a target of a bacteriophage 44AHJD product, preferably an ORF 12 or 25 product, e.g., as described for the previous aspect. Preferably the bacterium involved in the infection or risk of infection produces the identified target of the bacteriophage inhibitor protein or alternatively produces a homologous target compound. In preferred embodiments, the host organism is a plant or animal, preferably a mammal or bird, and more preferably, a human or other mammal described herein. Preferred embodiments include, without limitation, those as described for the preceding aspect.

Compounds useful for the methods of inhibiting, methods of treating, and pharmaceutical compositions can include novel compounds, but can also include compounds which had previously been identified for a purpose other than inhibition of bacteria. Such compounds can be utilized as described and can be included in pharmaceutical compositions.

By "treatment" or "treating" is meant administering a compound or pharmaceutical composition for prophylactic and/or therapeutic purposes. The term "prophylactic treatment" refers to treating a patient or animal that is not yet infected but is susceptible to or otherwise at risk of a bacterial infection. The term "therapeutic treatment" refers to administering treatment to a patient already suffering from infection.

The term "bacterial infection" refers to the invasion of the host organism, animal or plant, by pathogenic bacteria. This includes the excessive growth of bacteria which are normally present in or on the body of the organism, but more generally, a bacterial infection can be any situation in which the presence of a bacterial population(s) is damaging to a host organism. Thus, for example, an organism suffers from a bacterial infection when excessive numbers of a bacterial population are present in or on the organism's body, or when the effects of the presence of a bacterial population(s) is damaging to the cells, tissue, or organs of the organism.

The terms "administer", "administering", and "administration" refer to a method of giving a dosage of a compound or composition, e.g., an antibacterial pharmaceutical composition, to an organism. Where the organism is a mammal, the method is, e.g., topical, oral, intravenous, transdermal, intraperitoneal, intramuscular, or intrathecal. The preferred method of administration can vary depending on various factors, e.g., the components of the pharmaceutical composition, the site of the potential or actual bacterial infection, the bacterium involved, and the infection severity.

The term "mammal" has its usual biological meaning, referring to any organism of the Class Mammalia of higher vertebrates that nourish their young with milk secreted by mammary glands, e.g., mouse, rat, and, in particular, human, bovine, sheep, swine, dog, and cat.

In the context of treating a bacterial infection a "therapeutically effective amount" or "pharmaceutically effective amount" indicates an amount of an antibacterial agent, e.g., as disclosed for this invention, which has a therapeutic effect. This generally refers to the inhibition, to some extent, of the normal cellular functioning of bacterial cells that renders or contributes to bacterial infection.

The dose of antibacterial agent that is useful as a treatment is a "therapeutically effective amount." Thus, as used herein, a therapeutically effective amount means an amount of an antibacterial agent that produces the desired therapeutic effect as judged by clinical trial results and/or animal models. This amount can be routinely determined by one skilled in the art and will vary depending on several factors, such as the particular bacterial strain involved and the particular antibacterial agent used.

In connection with claims to methods of inhibiting bacteria and therapeutic or prophylactic treatments, "a compound active on a target of a bacteriophage inhibitor protein" or terms of equivalent meaning differ from administration of or contact with an intact phage naturally encoding the full-length inhibitor compound. While an intact phage may conceivably be incorporated in the present methods, the method at least includes the use of an active compound as specified different from a full length inhibitor protein naturally encoded by a bacteriophage and/or a delivery or contacting method different from administration of or contact with an intact phage naturally encoding the full-length protein. Similarly, pharmaceutical compositions described herein at least include an active compound or composition different from a phage naturally coding the full-length inhibitor protein, or such a full-length protein is provided in the composition in a form different from being encoded by an intact phage. Preferably the methods and compositions do not include an intact phage.

In accordance with the above aspects, the invention also provides antibacterial agents and compounds active on a bacterial target of a bacteriophage 44AHJD inhibitory, preferably ORF 12 or 25, where the target was uncharacterized as indicated above. As previously indicated, such active compounds include both novel compounds and compounds which had previously been identified for a purpose other than inhibition of bacteria. Such previously identified biologically active compounds can be used in embodiments of the above methods of inhibiting and treating. In preferred embodiments, the targets, bacteriophage, and active compound are as described herein for methods of inhibiting and methods of treating. Preferably the agent or compound is formulated in a pharmaceutical composition which includes a pharmaceutically acceptable carrier, excipient, or diluent. In addition, the invention provides agents, compounds, and pharmaceutical compositions where an active compound is active on an uncharacterized phage-specific site on the target.

In preferred embodiments, the target is as described for embodiments of aspects above.

Likewise, the invention provides a method of making an antibacterial agent. The method involves identifying a target of a bacteriophage 44AHJD ORF 12 or 25 product, screening a plurality of compounds to identify a compound active on the target, and

synthesizing the compound in an amount sufficient to provide a therapeutic effect when administered to an organism infected by a bacterium naturally producing the target.

In preferred embodiments, the identification of the target and identification of active compounds include steps or methods and/or components as described above (or otherwise herein) for such identification. Likewise, the active compound can be as described above, including fragments and derivatives of phage inhibitor proteins, peptidomimetics, and small molecules. As recognized by those skilled in the art, peptides can be synthesized by expression systems and purified, or can be synthesized artificially by methods well known in the art.

In the context of nucleic acid and/or amino acid sequences of this invention, the terms "correspond" and "corresponding" indicate that the sequence is at least 95% identical, preferably at least 97% identical, and more preferably at least 99% identical to a sequence from the specified phage genome or bacterial genome, a ribonucleotide equivalent, a degenerate equivalent (utilizing one or more degenerate codons), the translated product of a nucleic acid sequence, nucleic acid sequence(s) encoding for a specific polypeptide, or a homologous sequence, where the homolog provides functionally equivalent biological function. It is also understood that the terms "correspond" and "corresponding" indicate that a nucleic acid sequence corresponds to the polypeptide which corresponds to a protein encoded by a bacteriophage ORF, such as for example *S. aureus* bacteriophage 44AHJD, and a protein can correspond to the nucleic acid sequence(s) which encode therefor.

In embodiments where the bacterial target of a bacteriophage inhibitor ORF product, e.g., an inhibitory protein or polypeptide, the target is preferably encoded by a *S. aureus* nucleic acid coding sequence from a host bacterium for bacteriophage 44AHJD. Target sequences are described herein by reference to sequence source sites. The sequence encoding the target preferably corresponds to a *S. aureus* nucleic acid sequence available from numerous sources including *S. aureus* sequences deposited in GenBank, *S. aureus* sequences found in European Patent Application No. 97100110.7 to Human Genome Sciences, Inc. filed January 7, 1997, *S. aureus* sequences available from The Institute for Genome Research (TIGR) at internet address http://www., where the remainder of the address is tigr.org/tdb/mdb/mdbinprogress.html, *S. aureus* sequences available from the

Oklahoma University *S. aureus* sequencing project can be obtained by following directions provided on the internet address http://www., where the remainder of the address is .genome.ou.edu/staph.html, and *S. aureus* sequences available from internet address http://www., where the remainder of the address is .sanger.ac.uk/Projects/S_*aureus*/.

The amino acid sequence of a polypeptide target is readily provided by translating the corresponding coding region. For the sake of brevity, the sequences are not reproduced herein. Also, in preferred embodiments, a target sequence corresponds to a *S. aureus* coding sequences corresponding to a sequence listed in Table 6 herein. The listing in Table 6 describes *S. aureus* sequences currently deposited in GenBank. Again, for the sake of brevity, the sequences are described by reference to the GenBank entries instead of being written out in full herein. In cases where an entry for a coding region is not complete, the complete sequence can be readily obtained by routine methods, e.g., by isolating a clone in a phage 44AHJD host *S. aureus* genomic library, and sequencing the clone insert to provide the relevant coding region. The boundaries of the coding region can be identified by conventional sequence analysis and/or by expression in a bacterium in which the endogenous copy of the coding region has been inactivated and using subcloning to identify the functional start and stop codons for the coding region.

In an additional aspect, the present invention provides a nucleic acid segment which encodes a protein and corresponds to a segment of the nucleic acid sequence of an ORF (open reading frame) from *Staphylococcus aureus* bacteriophage 44AHJD corresponding to a sequence provided in Table 1. Preferably, the protein is a functional protein. One of ordinary skill in the art would recognize that bacteriophage possess genes which encode proteins which may be either beneficial or detrimental to a bacterial cell. Such proteins act to replicate DNA, translate RNA, manipulate DNA or RNA, and enable the phage to integrate into the bacterial genome. Proteins from bacteriophage can function as, for example, a polymerase, kinase, phosphatase, helicase, nuclease, topoisomerase, endonuclease, reverse transcriptase, endoribonuclease, dehydrogenase, gyrase, integrase, carboxypeptidase, proteinase, amidase, transcriptional regulators and the like, and/or the protein may be a functional protein such as a chaperon, capsid protein, head and tail proteins, a DNA or RNA binding protein, or a membrane protein, all of which are provided

as non-limiting examples. Proteins with functions such as these are useful as tools for the scientific community.

Thus, the present invention provides a group of novel proteins from bacteriophage which can be used as tools for biotechnical applications such as, for example, DNA and/or RNA sequencing, polymerase chain reaction and/or reverse transcriptase PCR, cloning experiments, cleavage of DNA and/or RNA, reporter assays and the like. Preferably, the protein is encoded by an open reading frame from the nucleic acid sequence of bacteriophage 44AHJD. Within the scope of the present invention are fragments of proteins and/or truncated portions of proteins which have been either engineered through automated protein synthesis, or prepared from nucleic acid segments which correspond to segments of the nucleic acid sequences of bacteriophage 44AHJD, and which are inserted into cells via plasmid vectors which can be induced to express the protein. It is understood by one of skill in the art that mutational analysis of proteins has been known to help provide proteins which are more stable and which have higher and/or more specific activities. Such mutations to proteins encoded by phage 44AHJD ORFs are also within the scope of the present invention, hence, the present invention also provides a mutated protein and/or the mutated nucleic acid segment from bacteriophage 44AHJD which encodes the protein.

In another aspect, the invention provides antibodies which bind proteins encoded by a nucleic acid segment which corresponds to the nucleic acid sequence of an ORF (open reading frame) from *Staphylococcus aureus* bacteriophage 44AHJD as provided in Table 1. Bacteriophages are bacterial viruses which contain nucleic acid sequences which encode proteins that can correspond to proteins of other bacteriophages and other viruses. Antibodies targeted for proteins encoded by nucleic acid segments of phage 44AHJD can serve to bind targets encoded by nucleic acid segments from other viruses which correspond to the sequences provided in Table 1. Furthermore, antibodies to proteins encoded by nucleic acid segments of phage 44AHJD can also bind to proteins from other viruses that share similar functions but may not share corresponding sequences. It is understood in the art that proteins with similar activities/functions from a variety of sources generally share motifs, regions, or domains which correspond. Thus, antibodies to motifs, regions, or domains of functional proteins from phage 44AHJD should be useful in detecting

corresponding proteins in other bacteriophages and viruses. Such antibodies can also be used to detect the presence of a virus sharing a similar protein. Preferably the virus to be detected is pathogenic to a mammal, such as a dog, cat, bovine, sheep, swine, or a human.

It has been determined that dnaN is a target for bacteriophage 44AHJD ORF 25 product, which acts as an inhibitory factor. Applicants have recognized the utility of the interaction in the development of antibacterial agents. Polypeptide and/or polynucleotide targets such as dnaN are critical targets for bacterial inhibition. *S. aureus* bacteriophage 44AHJD ORF 25 or derivatives or functional mimetics thereof are useful for inhibiting bacterial growth and the interaction, binding, inhibition and/or activation which occurs between polypeptides and/or polynucleotides, such as for example dnaN of *S. aureus* and 44AHJD ORF 25 may be used as a target for the screening and rational design of drugs or antibacterial agents. In addition to methods for directly inhibiting a target such as dnaN activity, methods of inhibiting a target such as dnaN expression are also attractive for antibacterial activity.

In a related aspect or in preferred embodiments, the present invention provides methods for identifying compounds which bind to, interact with, inhibit and/or activate an activity and/or expression of a polypeptide and/or polynucleotide of the invention, e.g., a polypeptide or polynucleotide that binds or interacts with a bacteriophage 44 AHJD inhibitory ORF, preferably ORF 12 or 25. Such methods comprise contacting a polypeptide and/or polynucleotide of the invention, such as for example a target or product of ORF 12 or 25 with a compound to be screened under conditions which permit binding or other interaction between the compound and the polypeptide and/or polynucleotide. The method, preferably allows assessment of the binding or other interaction with the compound being identified by associating the binding or interaction with a second component which is capable of providing a detectable signal in response to the binding or interaction of the polypeptide and/or polynucleotide with the compound. Determination of whether the compound binds to, interacts with, activates and/or inhibits an activity or expression of the polypeptide and/or polynucleotide is by detection of the presence or absence of a signal generated from the binding or interaction of the compound with the polypeptide and/or polynucleotide.

In preferred embodiments, the method involves the interaction of an ORF 12 or 25 product or fragment thereof with the corresponding bacterial target or fragment thereof that maintains the interaction with the ORF product or fragment. Interference with the interaction between the components can be monitored, and such interference is indicative of compounds that will inhibit, activate, or enhance the activity of the target molecule.

Preferably, compounds which are identified by methods of the present invention include, but are not limited to, small organic molecules, peptides, polypeptides and antibodies that bind to a polynucleotide and/or polypeptide of the invention, such as for example ORF 12 or 25 gene product or target thereof, and thereby inhibit or extinguish or enhance its activity or expression. Potential compounds also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a binding molecule, without inducing for example dnaN or dnaN homologues or peptido-mimetic derivatives, induced activities, thereby preventing the action or expression of *S. aureus* bacteriophage 44AHJD ORF 12 or 25 gene product or target thereof and/or for example dnaN polypeptides and/or polynucleotides by excluding *S. aureus* 44AHJD ORF 12 or 25 gene product or target thereof and/or for example dnaNpolypeptides and/or polynucleotides from binding.

Potential compounds also include small molecules that bind to and occupy the binding site of a polypeptide thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules. Other potential compounds include antisense molecules (see Okano, (1991) J. Neurochem. 56, 560; see also "Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression", CRC Press, Boca Raton, FL (1988), for a description of these molecules). Preferred potential compounds include compounds related to and variants of 44AHJD ORF 12 or 25 and of dnaN and any homologues and/or peptido-mimetics and/or fragments thereof. Other examples of potential polypeptide antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the polypeptide, e.g., a fragment of the ligands, substrates, receptors, enzymes, etc.; or

small molecules which bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See, e.g., Coligan et al., Current Protocols in Immunology 1(2): Chapter 5 (1991). Peptide modulators can also be selected by screening large random libraries of all possible peptides of a certain length.

Compounds derived from the polypeptide sequence of 44AHJD ORF 12 or 25 itself could represent fragments representing small overlapping peptide spanning the entire amino acid sequence of the protein. Fragments of 44AHJD ORF 12 or 25 can be produced by proteolytic digestion of the full-length protein as described above. Alternatively, suitable 44AHJD ORF 12 or 25 derived peptide or polypeptide fragments representative of the complete sequence of the protein can be chemically synthesized. For instance, in the multi-pin approach, peptides are simultaneously synthesized by the assembly of small quantities of peptides on plastic pins derivatized with an ester linker based on glycolate and 4-(hydroxymethyl) benzoate (Maeji et al. (1991) Pept Res, 4:142-6).

As used in the claims to describe the various inventive aspects and embodiments, "comprising" means including, but not limited to, whatever follows the word "comprising". Thus, use of the term "comprising" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present. By "consisting of" is meant including, and limited to, whatever follows the phrase "consisting of". Thus, the phrase "consisting of" indicates that the listed elements are required or mandatory and that no other elements may be present. By "consisting essentially of" is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase "consisting essentially of" indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not

be present depending upon whether or not they affect the activity or action of the listed elements.

Additional features and embodiments of the present invention will be apparent from the following Detailed Description and from the claims, all within the scope of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 is a flow schematic showing the manipulations necessary to convert pT0021, an arsenite inducible vector containing the luciferase gene, into pTHA or pTM, two ars inducible vectors, or pTMLac , a lactose-inducible promotor. a) Vector pTHA contains BamH I and Sal I cloning sites and a downstream HA epitope tag. b) and c)Vector pTM and pTMLac contain Bam HI and Hind III cloning sites and no HA epitope tag.

FIGURE 2 is a schematic representation of the cloning steps involved to place the DNA segments of any of ORFs 12, 25, or other sequences into vectors to assess inhibitory potential. a) For subcloning into pTHA, individual ORFs e.g. 44AHJD ORF 12 and 25 were amplified by the PCR using oligonucleotides targeting the start codon and the penultimated codon of the ORFs. Using this strategy, BamHI and SalI sites were positioned immediately upstream or downstream, respectively of these two codons. Following digestion with BamHI and SalI , the PCR fragments were subcloned into the same sites of pTHA.

b) For subcloning into pTM or pTMLac, (examplified for pTM in b) individual ORFs were amplified by the PCR using oligonucleotides targeting the ATG and stop codons of the ORFs. Using this strategy, Bam HI and Hind III sites were positioned immediately upstream or downstream, respectively of the start and stop codons of each ORF. Following digestion with Bam HI and Hind III, the PCR fragments were subcloned into the same sites of pTM or pTMLac. Clones were verified by PCR and direct sequencing.

FIGURE 3 shows a schematic representation of the functional assays used to characterize the bactericidal and bacteriostatic potential of all predicted ORFs (>33 amino

acids) encoded by bacteriophage 44AHJD. Fig. 3A) Functional assay on semi-solid support media. Fig. 3B) Functional assay in liquid culture.

FIGURE 4 shows the results of the functional assay on semi-solid support media to identify bacteriophage 44AHJD ORFs with anti-microbial activity. Figure 4 a) shows the lists of the 31 bacteriophage 44AHJD ORFs that were screened in the functional assay and Figure 4 b) shows inhibition of bacterial growth following induction of expression of phage 44AHJD ORF 12 and 25 from three clones of *Staphylococcus aureus* transformants tested at four different concentrations. One clone of *Staphylococcus aureus* transformed with the non-inhibitory ORF (77 bacteriophage ORF 30 cloned into pT vector) was used as control. From these experiments, it is clear that expression of these two ORFs leads to the inhibition of growth of *Staphylococcus aureus*.

FIGURE 5 A and B are the graphs of OD_{565} values and colony forming units (CFU) over time showing the results of functional assay in liquid media to assess bacteriostatic or bactericidal activity of bacteriophage 44AHJD ORF 12 and 25. Growth inhibition assays were performed as detailed in the Detailed Description. The OD_{565} values and the number of CFU were determined from cultures of *Staphylococcus aureus* transformants harboring a given bacteriophage inhibitory ORF, in the absence or presence of the inducer. The identity of the expression vector and subcloned ORF harbored by the *Staphylococcus aureus* is given at the top of the each graph. The value of OD and the number of CFU was also determined from non-induced and induced control cultures of *Staphylococcus aureus* transformants harboring a non-inhibitory phage ORF cloned into the same vector. Each graph represents the average obtained from three *Staphylococcus aureus* transformants.

FIGURE 6 shows the pattern of protein expression of the inhibitory ORF in *S. aureus* in the presence or in the absence of induction with sodium arsenite. In individual inhibitory ORF (44AHJD phage ORF 12 and 25) cloned into the pTHA vector, the HA tag is set inframe with the ORF and is positioned at the carboxy terminus of each ORF. An anti-HA tag antibody was used for the detection of the ORF expression. The identity of the

subcloned ORF harbored by the *Staphylococcus aureus* transformants is given at the top of the panel.

FIGURES 7A and 7B depict the results from affinity chromatography using GST and GST/44AHJD ORF 25 as ligands with a *S. aureus* extract prepared by French pressure cell lysis and sonication. Eluates from affinity columns containing the GST and GST/ORF25 ligands at 0, 0.1, 0.5, 1.0, and 2.0 mg/ml resin were resolved by SDS-12.5% PAGE. Proteins were visualized by silver staining. Micro-columns were eluted with: A) 1 M NaCl ABC (ACB; 20 mM Hepes pH 7.5, 10 % glycerol, 1 mM DTT, and 1 mM EDTA); and B) 1% SDS. Each molecular weight marker is approximately 100 ng. The lanes labeled ACB indicate eluates from a 2.0 mg/ml ligand column loaded only with ACB buffer containing 75 mM NaCl. The arrows indicate proteins specifically with GST/ORF25.

FIGURE 8 shows results of a tryptic peptide mass spectrum of the PT48 protein that interacted with 44AHJD ORF 25 and that was eluted with 1% SDS and labelled: PT48 in Figure 7B. The control band excised from the 48 kDA region of the gels containing PT48 did not contain PT48.

FIGURE 9 shows the identification of PT48 as *S. aureus* DNA-directed DNA polymerase III beta subunit protein from the Genbank database (accession number: 1084189).

FIGURE 10 shows the nucleotide and amino acid sequences of *S. aureus* DnaN.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Preliminarily the tables will be briefly described.

Table 1 shows the complete nucleotide sequence of the genome of *Staphylococcus aureus* bacteriophage 44AHJD.

Table 2 is a table from Alberts *et al.*, MOLECULAR BIOLOGY OF THE CELL 3rd ed., showing the redundancy of the "universal" genetic code.

Table 3 shows the nucleotide and predicted amino acid sequences of ORFs 12 and 25 from bacteriophage 44AHJD.

Table 4 Shows homology search results. BLAST analysis was performed with 44AHJD ORFs 12 and 25 against NCBI non-redundant nucleotide and Swissprot databases. The results of this search indicate that ORF 12 has significant homology to an hypothetical 15.7 Kd protein of Bacillus subtilis located in the SpoIIIC-CWLA intergenic region.

Table 5 shows the physiochemical parameters of phage 44AHJD ORFs 12 and 25. These include the primary amino acid sequence of the predicted protein, the average molecular weight, amino acid composition, theoretical pI and hydrophobicity properties (Kite-Doolittle scale).

Table 6 shows a list of *Staphylococcus aureus* sequences which may represent genes coding for target sequences for phage 44 AHJD ORFs 12 and 25 encoded antimicrobial proteins or peptides.

Table 7 shows genetic map and sequence position of the 73 orfs predicted to be encoded by phage 44AHJD that are greater than 33 amino acids.

Table 8 shows nucleotide and predicted amino acid sequence of all 73 orfs identified in phage 44AHJD.

The present invention is based on the identification of naturally-occurring DNA sequence elements encoding RNA or proteins with anti-microbial activity. Bacteriophages or phages, are viruses that infect and kill bacteria. They are natural enemies of bacteria and, over the course of evolution have perfected enzymes (products of DNA sequences) which enable them to infect a host bacteria, replicate their genetic material, usurp host metabolism, and ultimately kill their host. The scientific literature documents well the fact that many known bacteria have a large number of such bacteriophages that can infect and kill them (for example, see the ATCC bacteriophage collection at http://www.atcc.org) (Ackermann and DuBow, 1987). Although we know that many bacteriophages encode proteins which can significantly alter their host's metabolism, determination of the killing potential of a given

bacteriophage gene product can only be assessed by expressing the gene product in the target bacterial strain.

As indicated in the Summary above, the present invention is concerned with the use of bacteriophage 44AHJD coding sequences and the encoded polypeptides or RNA transcripts to identify bacterial targets for potential new antibacterial agents. Thus, the invention concerns the selection of relevant bacteria. Particularly relevant bacteria are those which are pathogens of a complex organism such as an animal, e.g., mammals, reptiles, and birds, and plants. However, the invention can be applied to any bacterium (whether pathogenic or not) for which bacteriophage are available or which are found to have cellular components closely homologous to components targeted by phage 44AHJD ORFs 12 and 25.

Identification of ORFs 12 and 25 and products from the phage which inhibit the host bacterium both provides an inhibitor compound and allows identification of the bacterial target affected by the phage-encoded inhibitor. Such a target is thus identified as a potential target for development of other antibacterial agents or inhibitors and the use of those targets to inhibit those bacteria. As indicated above, even if such a target is not initially identified in a particular bacterium, such a target can still be identified if a homologous target is identified in another bacterium. Usually, but not necessarily, such another bacterium would be a genetically closely related bacterium. Indeed, in some cases, an inhibitor encoded by phage 44AHJD ORF 12 or 25 can also inhibit such a homologous bacterial cellular component.

The demonstration that bacteriophage have adapted to inhibiting a host bacterium by acting on a particular cellular component or target provides a strong indication that that component is an appropriate target for developing and using antibacterial agents, e.g., in therapeutic treatments. Thus, the present invention provides additional guidance over mere identification of bacterial essential genes, as the present invention also provides an indication of accessibility of the target to an inhibitor, and an indication that the target is sufficiently stable over time (e.g., not subject to high rates of mutation) as phage acting on that target were able to develop and persist. Thus, the present invention identifies a

particular subset of essential cellular components which are particularly likely to be appropriate targets for development of antibacterial agents.

The invention also, therefore, concerns the development or identification of inhibitors of bacteria, in addition to the phage-encoded inhibitory proteins (or RNA transcripts), which are active on the targets of bacteriophage-encoded inhibitors. As described herein, such inhibitors can be of a variety of different types, but are preferably small molecules.

Target proteins of antibiotics generally provide a critical cell function such as DNA replication or cell wall biosynthesis. A proven approach in the discovery of a new drug is to obtain a target protein and to develop in vitro assays to interfere with the biological function of the protein. As described below for DNA polymerase III, biological machineries are often comprised of multiprotein complexes. Thus, any members of essential multiprotein complexes are hypothetical targets for drug development. However, the fact that a protein can be associated with certain biological function does not imply that it represents suitable intervention for new drugs development(Drews J. 2000, Science 287:1960-1964). For instance, although DNA replication is a well-known and essential process for bacterial growth, only a relatively small number of DNA replication complex proteins are targeted by antibiotics. In addition, screening of compounds that inhibit the function of a target must be selective. This underscores the necessity to develop new target-derived strategies that include the step of identification of the protein domain that should be specifically targeted for drug design.

DNA polymerase III holoenzyme is an essential component of bacterial DNA replication machinery. The holoenzyme contains seven different polypeptide chains. Some of these subunits are essential for normal DNA replication in vivo, as shown by the existence of temperature-sensitive (ts) mutations in genes encoding these polypeptides. Type III polymerases are exemplified by the replicase of the Gram-negative bacterium Escherichia coli, in which there are three separate components: a sliding clamp protein, a clamp loader complex and the DNA polymerase itself (Kelman et al. 1995, Annu. Rev. Biochem. 64: 171-200). The clamp loader is a multiprotein complex which uses ATP to assemble the sliding clamp around DNA. The DNA polymerase then binds to the sliding

clamp which tethers the polymerase to the DNA. As described in Biochemistry edited by Mathews and Holde [1995; the Benjamin/Cumming Publishing Company], the three subunits - alpha, epsilon and theta – form the polymerase core enzyme. The binding of the gamma complex, the clamp loader, converts the aggregate to a form refered to as DNA polymerase III star, polIII*. This binds to the beta subunit, the clamp slider, to produce the holoenzyme. The beta subunit is a homodimer and forms the ring shaped sliding clamp associated with DNA.

Although there are several studies on the mechanism of replication in Gram-negative bacteria there is little information about how Gram-positive organisms replicate their genetic material. Bacillus subtilis is the best characterized Gram-positive bacterium with respect to DNA replication (Barnes et al. 1995, Methods in Enzy. 262: 35-42), where many genes involved in DNA replication have been identified through the isolation of ts mutants. Studies in B. subtilis have identified a polymerase that appears to be involved in chromosome replication and is termed PolIII. The polC gene encodes Pol III, a large polypeptide likely corresponding to the alpha and epsilon subunits of the E. coli enzyme. B. subtilis and *Staphylococcus aureus* each also have a gene encoding a protein with 30% homology to the beta subunit of the E coli enzyme; however, neither protein has been purified or characterized (Alfonso and Fisher 1995, Mol. Gen. Gent. 246: 680-686). The *S. aureus* gene corresponding to the E. coli beta subunit is dnaN. *S. aureus* dnaN has been described in an International Patent Application entitled: "DNA REPLICATION PROTEINS AND THEIR USE TO SCREEN FOR CHEMICAL INHIBITORS" WO 99/37661.

The following description provides preferred methods for implementing the various aspects of the invention. However, as those skilled in the art will readily recognize, other approaches can be used to obtain and process relevant information. Thus, the invention is not limited to the specifically described methods. In addition, the following description provides a set of steps in a particular order. That series of steps describes the overall development involved in the present invention. However, it is clear that individual steps or portions of steps may be usefully practiced separately, and, further, that certain steps may be

performed in a different order or even bypassed if appropriate information is already available or is provided by other sources or methods.

Bacterial Targets for Antibiotics

The main promise for using a bacteriophage approach to drug discovery lies in the potential to determine the optimal molecular target. The target proteins of antibiotics generally provide a critical cell function such as DNA replication or cell wall biosynthesis. A proven approach in the discovery of a new drug is to obtain a target protein and to develop *in vitro* assays to interfere with the biological function of the protein. As described below for DNA polymerase III, biological machineries are often comprised of multiprotein complexes. Thus, any members of essential multiprotein complexes are hypothetical targets.

DNA polymerase III holoenzyme is an essential component of the DNA replication machinery. The holoenzyme contains seven different polypeptide chains. Some of these subunits are essential for normal DNA replication *in vivo*, as shown by the existence of temperature-sensitive (*ts*) mutations in genes encoding these polypeptides. Type III polymerases are exemplified by the replicase of the Gram-negative bacterium *Escherichia coli,* in which there are three separate components: a sliding clamp protein, a clamp loader complex and the DNA polymerase itself (Kelman et al., 1995, Annu. Rev. Biochem. 64: 171-200). As described in the literature, the three subunits, Alpha, epsilon and theta, form the polymerase core enzyme. The binding of the gamma complex, the clamp loader, converts the aggregate to a form refered to as DNA polymerase III star, *polIII**. This binds to the beta subunit, the clamp slider, to produce the holoenzyme. The beta subunit is a homodimer and forms the ring shaped sliding clamp associated with DNA.

Of the Gram-positive organisms, *Bacillus subtilis* and *Streptococcus pyogenes* are the best characterized with respect to DNA replication (Barnes et al., 1995, Methods in Enzy. 262:35-42, Bruck I. and O'Donnell, M. 2000, J.Bio0l.Chem. 275:28971-28983), where many genes involved in DNA replication have been identified through the isolation of *ts* mutants. Studies in *B.subtilis* have indentified a polymerase that appears to be involved in chromosome replication and is termed PolIII. The *polC* gene encodes Pol III, a large

polypeptide corresponding to the alpha and epsilon subunits of *E.coli* enzyme. *B. subtilis* and another Gram positive, *Staphylococcus aureus* each have a gene encoding a protein with 30% homology to the beta subunit of the *E.coli* enzyme. The *S. aureus* gene corresponding to the *E.coli* beta subunit is *dnaN*. *S. aureus* DnaN has been described in international patent application, "Dna Replication Proteins and Their Use To Screen for Chemical Inhibitors" WO 99/37661.

Identification of Inhibitory ORF

The methodology previously described in United States Provisional Application Pelletier, *et al.*, 60/168,777 filed December 1, 1999 was used to identify and characterize DNA sequences from *Staphylococcus aureus* bacteriophage 44AHJD that can act as anti-microbials. A nucleic acid segment isolated from *Staphylococcus aureus* bacteriophage 44AHJD encodes a protein, whose gene is referred to as ORF (open reading frame) 12 or 25. Thus, the present invention provides a nucleic acid sequence isolated from *Staphylococcus aureus* (Staph A or *S. aureus*) bacteriophage 44AHJD comprising at least a portion of the gene encoding ORF 12 or 25 with anti-microbial activity. The nucleic acid sequence can be isolated using a method similar to those described herein, or using another method. In addition, such a nucleic acid sequence can be chemically synthesized. Having the anti-microbial nucleic acid sequence of the present invention, parts thereof or oligonucleotides derived therefrom, other anti-microbial sequences from other bacteriophage sources using methods described herein or other methods can be isolated, including screening methods based on nucleic acid sequence hybridization.

The present invention provides the use of the Staph A bacteriophage 44AHJD anti-microbial DNA segment encoding ORF 12 or 25, as a pharmacological agent – either wholly or in part - as well as the use of peptidomimetics, developed from amino acid or nucleotide sequence knowledge of Staph A bacteriophage 44AHJD ORF 12 or 25. This can be achieved where the structure of the peptidomimetic compound corresponds to the structure of the active portion of ORF 12 or 25. In this analysis, the peptide backbone is transformed into a carbon-based hydrophobic structure that can retain cytostatic or cytocidal activity for the bacterium. This is done by standard medicinal chemistry methods, measuring growth

inhibition of the various molecules in liquid cultures or on solid medium. These mimetics also represent lead compounds for the development of novel antibiotics.

In this context, "corresponds" means that the peptidomimetic compound structure has sufficient similarities to the structure of the active portion of ORF 12 or 25 that the peptidomimetic will interact with the same molecule as the product of ORF 12 or 25 and preferably will elicit at least one cellular response in common which relates to the inhibition of the cell by the phage protein.

The invention also provides bacteriophage anti-microbial DNA segments from other phages based on nucleic acids and sequences hybridizing to the presently identified inhibitory ORF under high stringency conditions or sequences which are homologous as described above. The bacteriophage anti-microbial DNA segment from bacteriophage 44AHJD ORF 12 or 25 can be used to identify a related segment from another related or unrelated phage based on conditions of hybridization or sequence comparison.

The methodology previously described (U.S. Provisional Application 60/110,992, filed Dec. 3, 1998) is used to identify and characterize DNA sequences from *Staphylococcus* sp. bacteriophage 44 AHJD that can act as antimicrobials.

The *Staphylococcus aureus* propagating strain (PS 44A) was obtained from the Felix d'Herelle Reference Centre (#HER 1101) was used as a host to propagate its phage 44AHJD, also obtained from the Felix d'Herelle Reference Centre (#HER 101). We find that bacteriophage 44AHJD consists of 16,668 bp (Table 1) predicted to encode 73 ORFs greater than 33 amino acids (Table 7, Table 8). Computational analysis of the predicted protein products of *Staphylococcus aureus* bacteriophage 44AHJD, which detected homolgs in public databases, are listed in Table 6, along with the accompanying list of related proteins. protein products related to those deposited in public databases.

From this analysis, it is apparent that 3 genes (ORF 3, 7, and 8) are related to structural proteins found in other bacteriophages. These include genes predicted to encode a tail protein (ORF 3), an upper collar/connector protein of the phage virion (ORF 7), and a lower collar protein (ORF 8). Bioinformatics has also identified one gene whose product is likely involved in phage DNA synthesis. One gene (ORF 1) shows significant homology to DNA polymerases of a number of bacteriophages, bacteria and fungi, and the product of

this gene is likely responsible for replicating the genetic material of bacteriophage 44AHJD. ORF 2 encodes a protein with homology to the dinC gene of Bacillus subtilis which encodes a protein involved in teichoic acid biosynthesis. Teichoic acid is a polyphosphate polymer found in some, but not all, Gram positive organisms (and not in Gram negative organisms), where it is attached to the peptidoglycan layer. The phage protein may thus be involved in the synthesis of this material for incorporation into the cell wall, allowing enhanced lysis by the phage lysis enzymes or, as many enzymes can function in "reverse reactions", may be involved in its degradation allowing for penetration of the peptidoglycan and phage genome entry into the cell following adsorption. The similarity between *Staphylococcus aureus* bacteriophage 44AHJD and E. coli phage T7 indicates that they may share similar mechanisms of replication and growth. Both phages belong to the Pododviridae Family of bacteriophages and are members of the "T7-like" Genus of this Family (Ackermann and DuBow; VIth ICTV Report).

Two genes, ORF 9 and 12, were identified with the potential to encode antimicrobial protein products. The predicted product of ORF 9 is related to a class of genes which encodes lysozyme-like functions, enzymes which cleave linkages in the mucopolysaccharide cell wall structure of a variety of micro-organisms, including that from the *Staphylococcus aureus* bacteriophage. ORF 12 of *Staphylococcus aureus* bacteriophage 44AHJD shows homology to a set of lysis proteins from several bacteriophages. These lysis proteins are also referred to as holins, and represent phage encoded lysis functions required for transit of the phage murein hydrolases (lysozyme) to the periplasm, where it can digest the cell wall and thus lyse the bacterium.

Thus, the present invention seeks to provide a nucleic acid sequence isolated from *Staphylococcus aureus* bacteriophage 44AHJD comprising at least a portion of one of the genes described above with antimicrobial activity. For example, ORF 1 encodes a DNA polymerase function. It is possible that this polymerase utilizes host-derived accessory proteins for its activity when replicating the phage template, sequestering such proteins from use by the bacterial polymerase, resulting in inhibition of DNA replication, cell division, and cell growth. Alternatively, ORF 9 directly encodes a polypeptide with antimicrobial activity. ORF 9 is predicted to encode an amidase, a protein known to act as a cell wall

degrading enzyme. ORF 12 likely encodes a holin function required for transit of the phage amidase (gene 9 product) to the periplasm. When this type of gene product from Bacillus phage phi 29 (gene 14), was cloned in Escherichia coli, cell death ensued (Steiner et al., 1993). Thus, production of proteins from Bacillus phage phi 29 gene 14 in E. coli resulted in cell death, whereas production of protein from Bacillus phage phi 29 gene 14 concomitantly with the phi 29 lysozyme or unrelated murein-degrading enzymes led to lysis, suggesting that membrane-bound protein 14 induces a nonspecific lesion in the cytoplasmic membrane (Steiner et al., 1993).

The present invention also provides the use of the *Staphylococcus* bacteriophage 44 AHJD antimicrobial ORFs or ORF products as pharmacological agents, either wholly or in part and derivatives, as well as the use of corresponding peptidomimetics, developed from amino acid or nucleotide sequence knowledge derived from *Staphylococcus* bacteriophage 44 AHJD killer ORFs. This can be done where the structure of the peptidomimetic compound corresponds to the structure of the active portion of a product of an ORF. In this analysis, the peptide backbone is transformed into a carbon based hydrophobic structure that can retain cytostatic or cytocidal activity for the bacterium. This is done by standard medicinal chemistry methods, measuring growth inhibition of the various molecules in liquid cultures or on solid medium. These mimetics also represent lead compounds for the development of novel antibiotics. In this context, "corresponds" means that the peptidomimetic compound structure has sufficient similarities to the structure of the active portion of a product of one of the *Staphylococcus* ORFs listed in Table 7, that the peptidomimetic will interact with the same molecule as the product of the ORF, and preferably will elicit at least one cellular response in common which relates to the inhibition of the cell by the phage protein.

To validate the identity of an ORF as a killer ORF, it is preferably expressed in the host or other test bacterial organism and the effect of this expression on bacterial growth and replication is assessed. Therefore, all individual ORFs identified herein, e.g., those identified above, can be expressed, preferably overexpressed, in a suitable host bacterium e.g., a host *Staphylococcus* and the effect of this expression or overexpression on host metabolism and viability can be measured.

Individual ORFs can be resynthesized from the phage genomic DNA by the polymerase chain reaction (PCR) using oligonucleotide primers flanking the ORF on either side. Those skilled in the art are familiar with the design and synthesis of appropriate primer sequences. These single ORFs are preferably engineered so that they contain appropriate cloning sites at their extremities to allow their introduction into a new bacterial expression plasmid, allowing propagation in a standard bacterial host such as *E. coli*, but containing the necessary information for plasmid replication in the target microbe, *Staphylococcus* sp. (hereafter referred to as a shuttle vector). ORF nucleic acid sequences can also be provided by direct chemical synthesis based on the ORF sequences identified herein using conventional synthesis methods familiar to those skilled in the art.

This shuttle vector also preferably contains regulatory sequences that allow inducible expression of the introduced ORF. As the candidate ORF may encode a killer function that will eliminate the host, it is highly advantageous that it not be expressed (or at least not expressed at a substantial level) prior to testing for activity; thus screening for such sequences in a constitutive fashion is less likely to be successful (lethality). For example, regulatory sequences from the ars operon can be used to direct individual ORF expression in *Staphylococcus*. The ars operon encodes a series of proteins which normally mediate the extrusion of arsenite and several other trivalent oxyanions from the cells when they are exposed to such toxic substances in their environment. The operon encoding this detoxifying mechanism is normally silent and only induced when arsenite-related compounds are present.

Therefore, individual phage ORFs can be expressed in *Staphylococcus* or other suitable host in an inducible fashion by adding to the culture medium non-toxic arsenite concentrations during the growth of individual *Staphylococcus* (or other host cells) clones expressing such individual phage ORFs. Toxicity of the phage killer ORF for the host is monitored by reduction or arrest of growth under induction conditions, as measured by optical density in liquid culture or after plating the induced cultures on solid medium. Subsequently, interference of the phage ORF with the host biochemical pathways ultimately leading to reducing or arresting host metabolism can be measured by pulse chase

experiments using radiolabeled precursors of either DNA replication, RNA transcription, or protein synthesis.

Of course, other inducible regulatory sequences (e.g., promoters, operators, etc.) may be used (e.g., systems using positive induction of expression or systems using release of repression). A variety of such systems are known to those skilled in the art and can be utilized in the present invention.

Nucleic acid sequences of the present invention can be isolated using a method similar to those described herein or other methods known to those skilled in the art. In addition, such nucleic acid sequences can be chemically synthesized by well-known methods. Having the phage 44 AHJD ORFs, e.g., anti-bacterial ORFs of the present invention, portions thereof, or oligonucleotides derived therefrom as described, other antimicrobial sequences from other bacteriophage sources can be identified and isolated using methods described here or other methods, including methods utilizing nucleic acid hybridization and/or computer-based sequence alignment methods.

The invention also provides bacteriophage antimicrobial DNA segments from other phages based on nucleic acids and sequences hybridizing to the presently identified inhibitory ORF under high stringency conditions or sequences which are highly homologous. The bacteriophage antimicrobial DNA segment from bacteriophage 44 AHJD can be used to identify a related segment from another unrelated phage based on stringent conditions of hybridization or on being a homolog based on nucleic acid and/or amino acid sequence comparisons. As with the phage 44 AHJD inhibitory sequences, such homologous coding sequences and products can be used as antimicrobials, to construct active portions or derivatives, to construct peptidomimetics, and to identify bacterial targets.

Identification of Bacterial Targets

The present invention provides the use of *Staphylococcus* bacteriophage 44AHJD ORFs, such as for example ORFs 12 and 25 anti-microbial activity to identify essential host bacterium interacting proteins or other targets that could, in turn, be used for drug design

and/or screening of test compounds. Thus, the invention provides a method of screening for antibacterial agents by determining whether test compounds interact with (e.g., bind to) the bacterial target. The invention also provides a method of making an antibacterial agent based on production and purification of the protein or RNA product of bacteriophage 44AHJD ORF 12 or 25. The method involves identifying a bacterial target of the product of ORF 12 or 25, screening a plurality of compounds to identify a compound active on the target, and synthesizing the compound in an amount sufficient to provide a therapeutic effect when administered to an organism infected by a bacterium naturally producing the target. The rationale is that the product of ORFs 12 and 25 can physically interact and/or modify certain microbial host components to block their function.

A variety of methods are known to those skilled in the art for identifying interacting molecules and for identifying target cellular components. Several approaches and techniques are described below which can be used to identify the host bacterial pathway and protein that interact or are inhibited by ORF 12 or 25.

The first approach is a genetic screen for protein:protein interaction, e.g., either some form of two-hybrid screen or some form of suppressor screen. In one form of the two hybrid screen involving the yeast two hybrid system, the nucleic acid segment encoding ORF 12 or 25, or a portion thereof, is fused to the carboxyl terminus of the yeast Gal4 DNA binding domain to create a bait vector. A genomic DNA library of cloned *S. aureus* sequences which have been engineered into a plasmid where the *S. aureus* sequences are fused to the carboxyl terminus of the yeast GAL4 activation domain II (amino acids 768-881) is also generated. These plasmids are introduced alone, or in combination, into a yeast strain, e.g., AH109 (Clontech Laboratories, Palo Alto, CA), previously engineered with chromosomally integrated copies of the *E. coli* lacZ and the selectable HIS3 and ADE2 genes, both under Gal4 regulation (Durfee *et al.*, 1993). If the two proteins expressed in yeast interact, the resulting complex will activate transcription from promoters containing Gal4 binding sites. The *lacZ*, HIS3 and ADE2 genes, each driven by a promoter containing Gal4 binding sites, have been integrated into the genome of the host yeast system and are used for measuring protein-protein interactions. Such a system provides a physiological environment in which to detect potential protein interactions.

This system has been extensively used to identify novel protein-protein interaction partners and to map the sites required for interaction (for example, to identify interacting partners of translation factors (Qui *et al.*, 1998), transcription factors (Katagiri *et al.*, 1998), proteins involved in signal transduction (Endo *et al.*, 1997). Alternatively, a bacterial two-hybrid screen can be utilized to circumvent the need for the interacting proteins to be targeted to the nucleus, as is the case in the yeast system (Karimova *et al.*, 1998).

The protein targets of ORFs 12 and 25 can also be identified using bacterial genetic screens. One approach involves the overexpression of ORF 12 or 25 protein in mutagenized *S. aureus* followed by plating the cells and searching for colonies that can survive the anti-microbial activity of ORF 12 or 25. These colonies are then grown, their DNA extracted, and cloned into an expression vector that contains a replicon of a different incompatibility group from the plasmid expressing ORF 12 or 25. This library is then introduced into a wild-type Staph A bacterium in conjunction with an expression vector driving synthesis of ORF 12 or 25, followed by selection for surviving bacteria. Thus, Staph A DNA fragments from the survivors presumably contain a DNA fragment from the original mutagenized Staph A genome that can protect the cell from the antimicrobial activity of ORF 12 or 25. This fragment can be sequenced and compared with that of the bacterial host to determine in which gene the mutation lies. This approach enables one to determine the targets and pathways that are affected by the killing function.

Alternatively, the bacterial targets can be determined in the absence of selecting for mutations using the approach known as "multicopy suppression". In this approach, the DNA from the wild type Staph A host is cloned into an expression vector that can coexist with the one containing ORF 12 or 25. Those plasmids that contain host DNA fragments and genes which protect the host from the anti microbial activity of ORF 12 or 25 can then be isolated and sequenced to identify putative targets and pathways in the host bacteria.

Another approach is based on identifying protein:protein interactions between the product of ORF 12 or 25 and *S. aureus* host proteins, using a biochemical approach based on affinity chromatography. This approach has been used to identify interactions between lambda phage proteins and proteins from their *E. coli* host (Sopta *et al.*, 1995). The product of ORF 12 or 25 is fused to a tag (e.g. -glutathione-S-transferase) after insertion in a

commercially available plasmid vector which directs high-level expression after induction of the responsive promoter driving the fusion protein. The fusion protein is expressed in *E. coli*, purified, and immobilized on a solid phase matrix. Total cell extracts from *S. aureus* are then passed through the affinity matrix containing the immobilized phage ORF fusion protein; host proteins retained on the column are then eluted under different conditions of ionic strength, pH, and detergents and identified by gel electrophoresis. They are recovered from the gel by transfer to a high affinity membrane. The proteins are individually digested to completion with a protease (e.g.-trypsin) and either molecular mass or the amino acid sequence of the tryptic fragments can be determined by mass spectrometry using MALDI-TOF technology (Qin *et al.*, 1997). The sequence of the individual peptides from a single protein are then analyzed by a bioinformatics approach to identify the *S. aureus* protein interacting with the phage ORF. This is performed by a computer search of the *S. aureus* genome for the identified sequence. Alternatively, tryptic peptide fragments of the *S. aureus* genome can be predicted by computer software based on the nucleotide sequence of the genome, and the predicted molecular mass of peptide fragments generated in silico compared to the molecular mass of the peptides obtained from each interacting protein eluted from the affinity matrix.

In addition, an oligonucleotide cocktail can be synthesized based on the primary amino acid sequence determined for an interacting *S. aureus* protein fragment. This oligonucleotide cocktail would comprise a mixture of oligonucleotides based on the nucleotide sequences of the primary amino acid of the predicted peptide, but in which all possible codons for a particular amino acid sequence are present in a subset of the oligonucleotide pool. This cocktail can then be used as a degenerate probe set to screen, by hybridization to genomic or cDNA libraries, to isolate the corresponding gene.

Alternatively, antibodies raised to peptides which correspond to an interacting *S. aureus* protein fragment can be used to screen expression libraries (genomic or cDNA) to identify the gene encoding the interacting protein.

Vectors

The invention also provides vectors, preferably expression vectors, harboring the anti-microbial DNA nucleic acid segment of the invention in an expressible form, and cells transformed with the same. Such cells can serve a variety of purposes, such as in vitro models for the function of the anti-microbial nucleic acid segment and screening for downstream targets of the anti-microbial nucleic acid segment, as well as expression to provide relatively large quantities of the inhibitory product.

Thus, an expression vector harboring the anti-microbial nucleic acid segment or parts thereof (Staph A bacteriophage 44AHJD ORF 12 or 25) can also be used to obtain substantially pure protein. Well-known vectors, such as the pGEX series (available from Pharmacia), can be used to obtain large amounts of the protein which can then be purified by standard biochemical methods based on charge, molecular mass, solubility, or affinity selection of the protein by using gene fusion techniques (such as GST fusion, which permits the purification of the protein of interest on a glutathione column). Other types of purification methods or fusion proteins could also be used as recognized by those skilled in the art.

Likewise, vectors containing bacteriophage 44AHJD ORFs 12 and 25 can be used in methods for identifying targets of the encoded antibacterial ORF product, e.g., as described above, and/or for testing inhibition of homologous bacterial targets or other potential targets in bacterial species other than *Staphylococcus aureus*.

Antibodies

Antibodies, both polyclonal and monoclonal, can be prepared against the protein encoded by a bacteriophage anti-microbial DNA segment of the invention (e.g., Staph A bacteriophage 44AHJD ORF 12 or 25) by methods well known in the art. Protein for preparation of such antibodies can be prepared by purification, usually from a recombinant cell expressing the specified ORF or fragment thereof. Those skilled in the art are familiar with methods for preparing polyclonal or monoclonal antibodies (See, e.g., Antibodies: A Laboratory Manual, Harlow and Lane, Cold Spring Harbor Laboratory, CSHL Press, N.Y., 1988).

Such antibodies can be used for a variety of purposes including affinity purification of the protein encoded by the bacteriophage anti-microbial DNA segment, tethering of the protein encoded by the bacteriophage anti-microbial DNA segment to a solid matrix for purposes of identifying interacting host bacterium proteins, and for monitoring of expression of the protein encoded by the bacteriophage anti-microbial DNA segment.

Recombinant Cells

Bacterial cells containing an inducible vector regulating expression of the bacteriophage anti-microbial DNA segment can be used to generate an animal model system for the study of infection by the host bacterium. The functional activity of the proteins encoded by the bacteriophage anti-microbial DNA segments, whether native or mutated, can be tested in animal in vitro or in vivo models.

While such cells containing inducible expression vectors is preferred, other recombinant cells containing a recombinant phage 44AHJD ORF 12 or 25 sequence or portion thereof are also provided by the present invention.

Also, a recombinant cell may contain a recombinant sequence encoding at least a portion of a protein which is a target of phage 44AHJD ORF 12 or 25 inhibitory ORF product.

In the context of this invention, in connection with nucleic acid sequences, the term "recombinant" refers to nucleic acid sequences which have been placed in a genetic location by intervention using molecular biology techniques, and does not include the relocation of phage sequences during or as a result of phage infection of a bacterium or normal genetic exchange processes such as bacterial conjugation.

Derivatization of identified anti-microbials

In cases where the identified anti-microbials above are peptidic compounds, the in vivo effectiveness of such compounds may be advantageously enhanced by chemical modification using the natural polypeptide as a starting point and incorporating changes that provide advantages for use, for example, increased stability to proteolytic degradation, reduced antigenicity, improved tissue penetration, and/or improved delivery characteristics.

In addition to active modifications and derivative creations, it can also be useful to provide inactive modifications or derivatives for use as negative controls or introduction of immunologic tolerance. For example, a biologically inactive derivative which has essentially the same epitopes as the corresponding natural antimicrobial can be used to induce immunological tolerance in a patient being treated. The induction of tolerance can then allow uninterrupted treatment with the active anti-microbial to continue for a significantly longer period of time.

Modified anti-microbial polypeptides and derivatives can be produced using a number of different types of modifications to the amino acid chain. Many such methods are known to those skilled in the art. The changes can include, for example, reduction of the size of the molecule, and/or the modification of the amino acid sequence of the molecule. In addition, a variety of different chemical modifications of the naturally occurring polypeptide can be used, either with or without modifications to the amino acid sequence or size of the molecule. Such chemical modifications can, for example, include the incorporation of modified or non-natural amino acids or non-amino acid moieties during synthesis of the peptide chain, or the post-synthesis modification of incorporated chain moieties.

The oligopeptides of this invention can be synthesized chemically or through an appropriate gene expression system. Synthetic peptides can include both naturally occurring amino acids and laboratory synthesized, modified amino acids.

Also provided herein are functional derivatives of anti-microbial proteins or polypeptides. By "functional derivative" is meant a "chemical derivative," "fragment," "variant," "chimera," or "hybrid" of the polypeptide or protein, which terms are defined below. A functional derivative retains at least a portion of the function of the protein, for example, reactivity with a specific antibody, enzymatic activity or binding activity.

A "chemical derivative" of the complex contains additional chemical moieties not normally a part of the protein or peptide. Such moieties may improve the molecule's solubility, absorption, biological half-life, and the like. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, and the like. Moieties capable of mediating such effects are disclosed in Genaro 1995, Remington's Pharmaceutical Science. Procedures for coupling such moieties

to a molecule are well known in the art. Covalent modifications of the protein or peptides are included within the scope of this invention. Such modifications may be introduced into the molecule by reacting targeted amino acid residues of the peptide with an organic derivatizing agent that is capable of reacting with selected side chains or terminal residues, as described below.

Cysteinyl residues most commonly are reacted with alpha-haloacetates (and corresponding amines), such as chloroacetic acid or chloroacetamide, to give carboxymethyl or carboxyamidomethyl derivatives. Cysteinyl residues also are derivatized by reaction with bromotrifluoroacetone, chloroacetyl phosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide, methyl 2-pyridyl disulfide, p-chloromercuribenzoate, 2-chloromercuri-4-nitrophenol, or chloro-7-nitrobenzo-2-oxa-1,3-diazole.

Histidyl residues are derivatized by reaction with diethylprocarbonate at pH 5.5-7.0 because this agent is relatively specific for the histidyl side chain. Para-bromophenacyl bromide also is useful; the reaction is preferably performed in 0.1 M sodium cacodylate at pH 6.0.

Lysinyl and amino terminal residues are reacted with succinic or other carboxylic acid anhydrides. Derivatization with these agents has the effect of reversing the charge of the lysinyl residues. Other suitable reagents for derivatizing primary amine- containing residues include imidoesters such as methyl picolinimidate; pyridoxal phosphate; pyridoxal; chloroborohydride; trinitrobenzenesulfonic acid; O-methylisourea; 2,4 pentanedione; and transaminase-catalyzed reaction with glyoxylate.

Arginyl residues are modified by reaction with one or several conventional reagents, among them phenylglyoxal, 2,3-butanedione, 1,2-cyclohexanedione, and ninhydrin. Derivatization of arginine residues requires that the reaction be performed in alkaline conditions because of the high pKa of the guanidine functional group. Furthermore, these reagents may react with the groups of lysine as well as the arginine alpha-amino group.

Tyrosyl residues are well-known targets of modification for introduction of spectral labels by reaction with aromatic diazonium compounds or tetranitromethane. Most commonly, N-acetylimidizol and tetranitromethane are used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively.

Carboxyl side groups (aspartyl or glutamyl) are selectively modified by reaction carbodiimide (R'-N-C-N-R') such as 1-cyclohexyl-3-(2-morpholinyl(4-ethyl) carbodiimide or 1-ethyl-3-(4-azonia-4,4-dimethylpentyl) carbodiimide. Furthermore, aspartyl and glutamyl residues are converted to asparaginyl and glutaminyl residues by reaction with ammonium ions.

Glutaminyl and asparaginyl residues are frequently deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues falls within the scope of this invention.

Derivatization with bifunctional agents is useful, for example, for cross-linking component peptides to each other or the complex to a water-insoluble support matrix or to other macromolecular carriers. Commonly used cross-linking agents include, for example, 1,1-bis (diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), and bifunctional maleimides such as bis-N-maleimido-1,8-octane. Derivatizing agents such as methyl-3-[p-azidophenyl) dithiolpropioimidate yield photoactivatable intermediates that are capable of forming crosslinks in the presence of light. Alternatively, reactive water-insoluble matrices such as cyanogen bromide-activated carbohydrates and the reactive substrates described in U.S. Patent Nos. 3,969,287; 3,691,016; 4,195,128; 4,247,642; 4,229,537; and 4,330,440 are employed for protein immobilization.

Other modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains (Creighton, T.E., Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)), acetylation of the N-terminal amine, and, in some instances, amidation of the C-terminal carboxyl groups.

Such derivatized moieties may improve the stability, solubility, absorption, biological half-life, and the like. The moieties may alternatively eliminate or attenuate any undesirable side effect of the protein complex. Moieties capable of mediating such effects are disclosed, for example, in Genaro 1995, Remington's Pharmaceutical Science.

The term "fragment" is used to indicate a polypeptide derived from the amino acid sequence of the protein or polypeptide having a length less than the full-length polypeptide from which it has been derived. Such a fragment may, for example, be produced by proteolytic cleavage of the full-length protein. Preferably, the fragment is obtained recombinantly by appropriately modifying the DNA sequence encoding the proteins to delete one or more amino acids at one or more sites of the C-terminus, N-terminus, and/or within the native sequence.

Another functional derivative intended to be within the scope of the present invention is a "variant" polypeptide which either lacks one or more amino acids or contains additional or substituted amino acids relative to the native polypeptide. The variant may be derived from a naturally occurring polypeptide by appropriately modifying the protein DNA coding sequence to add, remove, and/or to modify codons for one or more amino acids at one or more sites of the C-terminus, N-terminus, and/or within the native sequence.

A functional derivative of a protein or polypeptide with deleted, inserted and/or substituted amino acid residues may be prepared using standard techniques well-known to those of ordinary skill in the art. For example, the modified components of the functional derivatives may be produced using site-directed mutagenesis techniques (as exemplified by Adelman *et al.*, 1983, DNA 2:183; Sambrook *et al.*, 1989) wherein nucleotides in the DNA coding sequence are modified such that a modified coding sequence is produced, and thereafter expressing this recombinant DNA in a prokaryotic or eukaryotic host cell, using techniques such as those described above. Alternatively, components of functional derivatives of complexes with amino acid deletions, insertions and/or substitutions may be conveniently prepared by direct chemical synthesis, using methods well-known in the art.

Insofar as other anti-microbial inhibitor compounds identified by the invention described herein may not be peptidal in nature, other chemical techniques exist to allow their suitable modification, as well, and according the desirable principles discussed above.

Administration and Pharmaceutical Compositions

For the therapeutic and prophylactic treatment of infection, the preferred method of preparation or administration of anti-microbial compounds will generally vary depending on

the precise identity and nature of the anti-microbial being delivered. Thus, those skilled in the art will understand that administration methods known in the art will also be appropriate for the compounds of this invention. Pharmaceutical compositions are prepared, as understood by those skilled in the art, to be appropriate for therapeutic use. Thus, generally the components and composition are prepared to be sterile and free of components or contaminants which would pose an unacceptable risk to a patient. For compositions to be administered internally is is generally important that the composition be pyrogen free, for example.

The particularly desired anti-microbial can be administered to a patient either by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s). In treating an infection, a therapeutically effective amount of an agent or agents is administered. A therapeutically effective dose refers to that amount of the compound that results in amelioration of one or more symptoms of bacterial infection and/or a prolongation of patient survival or patient comfort.

Toxicity, therapeutic and prophylactic efficacy of anti-microbials can be determined by standard pharmaceutical procedures in cell cultures and/or experimental organisms such as animals, e.g., for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized.

For any compound identified and used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. Such information can be used to more accurately determine useful doses in organisms such as plants and animals, preferably mammals, and most preferably humans. Levels in plasma

may be measured, for example, by HPLC or other means appropriate for detection of the particular compound.

The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition (see e.g. Fingl et. al., in The Pharmacological Basis of Therapeutics, 1975, Ch. 1 p.1).

It should be noted that the attending physician would know how and when to terminate, interrupt, or adjust administration due to toxicity, organ dysfunction, or other systemic malady. Conversely, the attending physician would also know to adjust treatment to higher levels if the clinical response were not adequate (precluding toxicity). The magnitude of an administered dose in the management of the disorder of interest will vary with the severity of the condition to be treated and the route of administration. The severity of the condition may, for example, be evaluated, in part, by standard prognostic evaluation methods. Further, the dose and perhaps dose frequency, will also vary according to the age, body weight, and response of the individual patient. A program comparable to that discussed above also may be used in veterinary or phyto medicine.

Depending on the specific infection target being treated and the method selected, such agents may be formulated and administered systemically or locally, i.e., topically. Techniques for formulation and administration may be found in Genaro 1995, Remington's Pharmaceutical Science. Suitable routes may include , for example, oral, rectal, transdermal, vaginal, transmucosal, intestinal, parenteral, intramuscular, subcutaneous, or intramedullary injections, as well as intrathecal, intravenous, or intraperitoneal injections.

For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

Use of pharmaceutically acceptable carriers to formulate identified anti-microbials of the present invention into dosages suitable for systemic administration is within the scope of the invention. With proper choice of carrier and suitable manufacturing practice, the compositions of the present invention, in particular those formulated as solutions, may be

administered parenterally, such as by intravenous injection. Appropriate compounds can be formulated readily using pharmaceutically acceptable carriers well known in the art into dosages suitable for oral administration. Such carriers enable the compounds of the invention to be formulated as tablets, pills, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated.

Agents intended to be administered intracellularly may be administered using techniques well known to those of ordinary skill in the art. For example, such agents may be encapsulated into liposomes, then administered as described above. Liposomes are spherical lipid bilayers with aqueous interiors. All molecules present in an aqueous solution at the time of liposome formation are incorporated into the aqueous interior. The liposomal contents are both protected from the external microenvironment and, because liposomes fuse with cell membranes, are efficiently delivered into the cell cytoplasm. Additionally, due to their hydrophobicity, small organic molecules may be directly administered intracellularly.

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. Determination of the effective amounts is well within the capability of those skilled in the art.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. The preparations formulated for oral administration may be in the form of tablets, dragees, capsules, or solutions, including those formulated for delayed release or only to be released when the pharmaceutical reaches the small or large intestine.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levitating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active anti-microbial compounds in water-soluble form. Alternatively, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid

esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Pharmaceutical preparations for oral use can be obtained by combining the active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added.

The above methodologies may be employed either actively or prophylactically against an infection of interest.

To identify DNA segments of Staph A bacteriophage 44AHJD capable of acting as anti-microbial agents, a strategy described in United States Provisional Application Pelletier, *et al.*, 60/168,777 filed December 1, 1999 was employed. In essence, the procedure involved sequence characterization of the bacteriophage, identification of protein coding regions (open reading frames or ORFs), subcloning of all ORFs into an appropriate inducible expression vector, transfer of the ORF subclones into Staph. A, followed by induction of ORF expression and assessment of effect on growth. We employed discovery steps as described in the Examples.

There are a number of methods provided for determining if a compound binds to, interacts with, activates or inhibits an activity or expression of a polypeptide and/or polynucleotide target such as dnaN for example. Screening methods that measure the binding of a screened compound to an ORF 12 or 25 target or product, such as for example dnaN polypeptide and/or polynucleotide, or binding of a screened compound to cells or supports bearing an ORF 12 or 25 target or product polypeptide or a fusion protein comprising the target, by direct labeling or indirectly associating a label with a screened compound are within the scope of the present invention. A screening method of the invention may involve competition for binding of a labeled competitor binding molecule, polypeptide and/or polynucleotide, such as bacteriophage 44AHJD ORF 12 or 25 target or product, or a fragment which binds to a target protein such as dnaN.

Accordingly, the present invention provides methods of screening compounds to identify those compounds which modulate, bind to, interact with, inhibit and/or activate activity or expression of a polypeptide or polynucleotide of the invention.

Example I. Growth of Staph A bacteriophage 44AHJD and purification of genomic DNA.

The *Staphylococcus aureus* propagating strain (PS 44A) (Felix d'Herelle Reference Centre #HER 1101, Ottawa, Canada) was used as a host to propagate its respective phage 44AHJD (Felix d'Herelle Reference Centre #HER101). Two rounds of plaque purification of phage 44AHJD were performed on soft agar essentially as described in Sambrook et al (1989). Briefly, the PS 44AHJD strain was grown overnight at 37°C in Nutrient broth [NB:

0.3% Bacto beef extract, 0.5% Bacto peptone (Difco Laboratories) and 0.5% NaCl (w/v)].The culture was then diluted 20x in NB and incubated at 37°C until the OD_{540}= .2 (early log phase) with constant agitation. In order to obtain single plaques, phage 44AHJD was subjected to 10-fold serial dilutions using phage buffer (1 mM $MgSO_4$, 5 mM $MgCl_2$, 80 mM NaCl and 0.1% Gelatin (w/v)) and 10 µl of each dilution was used to infect 0.5 ml of the cell suspension in the presence of 400 µg/ml $CaCl_2$. After incubation of 15 min at room temperature (RT), 2 ml of melted soft agar kept at 45°C (NB supplemented with 0.6% agar) was added to the mixture and poured onto the surface of 100 mm nutrient agar plates (0.3% Bacto Beef extract, 0.5% Bacto peptone, 0.5% NaCl and 1.5% Bacto agar (w/v)). After overnight incubation at 30°C, a single plaque was isolated, resuspended in 1 ml of phage buffer by end over end rotation for 2 hrs at 20°C, and the phage suspension was diluted and used for a second infection as described above. After overnight incubation at 30°C, a single plaque was isolated and used as a stock.

The propagation procedure for bacteriophage 44AHJD was modified from the agar layer method of Swanstörm and Adams (1951). Briefly, the PS 44A strain was grown to stationary phase overnight at 37°C in Nutrient broth. The culture was then diluted twenty-fold in NB and incubated at 37°C until the OD_{540}= .2. The suspension ($15x10^7$ Bacteria) was then mixed with $15x10^5$ plaque forming units (pfu) to give a ratio of 100-bacteria/phage particle in the presence of 400 µg/ml of $CaCl_2$. After incubation for 15 min at 20°C, 7.5 ml of melted soft agar (NB plus 0.6% agar) were added to the mixture and poured onto the surface of 150 mm nutrient agar plates and incubated 16 hrs at 37°C. To collect the phage plate lysate, 20 ml of NB were added to each plate and the soft agar layer was collected by scrapping off with a clean microscope slide followed by shaking of the agar suspension for 5 min to break up the agar. The mixture was then centrifuged for 10 min at 4,000 RPM (2,830 x g) in a JA-10 rotor (Beckman) and the supernatant fluid (lysate) was collected and subjected to a treatment with 10 µg /ml of DNase I and RNase A for 30 min at 37°C. To precipitate the phage particles, the phage suspension was adjusted to 10% (w/v) PEG 8000 and 0.5 M of NaCl followed by incubation at 4°C for 16 hrs. The phage was recovered by centrifugation at 4,000 rpm (3,500 x g) for 20 min at 4°C on a GS-6R table top centrifuge (Beckman). The pellet was resuspended with 2 ml of phage buffer (1 mM $MgSO_4$, 5 mM

$MgCl_2$, 80 mM NaCl and 0.1% Gelatin). The phage suspension was extracted with 1 volume of chloroform and further purified by centrifugation on a cesium chloride step gradient as described in Sambrook *et al.* (1989), using a TLS 55 rotor centrifuged in an Optima TLX ultracentrifuge (Beckman) for 2 hr at 28,000 rpm (67,000 x g) at 4°C. Banded phage was collected and ultracentrifuged again on an isopycnic cesium chloride gradient (1.45 g/ml) at 40,000 rpm (64,000 x g) for 24 h at 4°C using a TLV rotor (Beckman). The phage was harvested and dialyzed for 4 h at room temperature against 4 L of dialysis buffer consisting of 10 mM NaCl, 50 mM Tris-HCl [pH 8] and 10 mM $MgCl_2$. Phage DNA was prepared from the phage suspension by adding 20 mM EDTA, 50 ug/ml Proteinase K and 0.5% SDS and incubating for 1 hr at 65°C, followed by successive extractions with 1 volume of phenol, 1 volume of phenol-chloroform and 1 volume of chloroform. The DNA was then dialyzed overnight at 4°C against 4 L of TE (10 mM Tris HCl [pH 8.0], 1mM EDTA).

Example II. DNA sequencing of Bacteriophage 44AHJD genome

Four micrograms of phage DNA was diluted in 200 µl of TE (10 mM Tris, [pH 8.0], 1 mM EDTA) in a 1.5 ml eppendorf tube and sonication was performed (550 Sonic Dismembrator™, Fisher Scientific). Samples were sonicated under an amplitude of 3 µm with bursts of 5 s spaced by 15 s cooling in ice/water for 3 to 4 cycles. The sonicated DNA was then size fractionated by electrophoresis on 1% agarose gels utilizing TAE (1 x TAE is: 40 mM Tris-acetate, 1 mM EDTA [pH 8.0]) as the running buffer. Fractions ranging from 1 to 2 kbp were excised from the agarose gel and purified using a commercial DNA extraction system according to the instructions of the manufacturer (Qiagen), with a final elution of 50 µl of 1 mM Tris HCl [pH 8.5].

The ends of the sonicated DNA fragments were repaired with a combination of T4 DNA polymerase and the Klenow fragment of *E. coli* DNA polymerase I, as follows. Reactions were performed in a reaction mixture (final volume, 100 µl) containing sonicated phage DNA, 10 mM Tris-HCl [pH 8.0], 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 50 µg/ml BSA, 100 µM of each dNTP and 15 units of T4 DNA polymerase (New England Biolabs) for 20 min at 12°C followed by addition of 12.5 units of Klenow large fragment (New England Biolabs) for 15 min at room temperature. The reaction was stopped by two

phenol/chloroform extractions and the DNA was precipitated with ethanol and the final DNA pellet was resuspended in 20 µl of H2O.

Blunt-ended DNA fragments were cloned by ligation directly into the Hinc II site of pKSII+ vector (Stratagene) dephosphorylated by treatment with calf intestinal alkaline phosphatase (New England Biolabs). A typical ligation reaction contained 100 ng of vector DNA, 2 to 5 µl of repaired sonicated phage DNA (50-100 ng) in a final volume of 20 µl containing 800 units of T4 DNA ligase (New England Biolabs) and was incubated overnight at 16°C. Transformation and selection of bacterial clones containing recombinant plasmids was performed in *E. coli* DH10β according to standard procedures (Sambrook *et al.*, 1989).

Recombinant clones were picked from agar plates into 96-well plates containing 100 µl LB and 100 µg/ml ampicillin and incubated at 37°C. The presence of phage DNA insert was confirmed by PCR amplification using T3 and T7 primers flanking the Hinc II cloning site of the pKS II+ vector. PCR amplification of foreign insert was performed in a 15 µl reaction volume containing 10 mM Tris HCl [pH 8.3], 50 mM KCl, 1.5 mM $MgCl_2$, 0.02% gelatin, 1 µM primer, 187.5 µM each dNTP, and 0.75 units Taq polymerase (BRL). The thermocycling parameters were as follows: 2 min initial denaturation at 94°C for 2 min, followed by 20 cycles of 30 sec denaturation at 94°C, 30 sec annealing at 57°C, and 2 min extension at 72°C, followed by a single extension step at 72°C for 10 min. Clones with insert sizes of 1 to 2 kbp were selected and plasmid DNA was prepared from the selected clones using QIAprep™ spin miniprep kit (Qiagen).

The nucleotide sequence of the extremities of each recombinant clone was determined using an ABI 377-36 automated sequencer with two types of chemistry: ABI prism Big Dye™ primer cycle sequencing (21M13 primer: #403055)(M13REV primer: #403056) or ABI prism Big Dye™ terminator cycle sequencing ready reaction kit (Applied Biosystems, #4303152). To ensure co-linearity of the sequence data and the genome, all regions of phage genome were sequenced at least once from both directions on two separate clones. In areas that this criteria was not initially met, a sequencing primer was selected and phage DNA was used directly as sequencing template employing ABI prism Big Dye™ terminator cycle sequencing ready reaction kit.

Example III. Bioinformatic management of primary nucleotide sequence.

Sequence contigs were assembled using Sequencher™ 3.1 software (GeneCodes). To close contig gaps, sequencing primers were selected near the edge of the contigs. Phage DNA was used directly as sequencing template employing ABI prism BIG DYE™ terminator cycle sequencing ready reaction kit. The complete sequence of bacteriophage 44AHJD is shown in Table 1.

A software program was developed and used on the assembled sequence of bacteriophage 44AHJD to identify all putative ORFs larger than 33 codons. Other ORF identification software can also be utilized, preferably programs which allow alternative start codons. The software scans the primary nucleotide sequence starting at nucleotide #1 for an appropriate start codon. Three possible selections can be made for defining the nature of the start codon; I) selection of ATG, II) selection of ATG or GTG, and III) selection of either ATG, GTG, TTG, CTG, ATT, ATC, and ATA. This latter initiation codon set corresponds to the one reported by the NCBI (http://www.ncbi.nlm.nih.gov/htbin-post/Taxonomy/wprintgc?mode=c) for the bacterial genetic code.

When an appropriate start codon is encountered, a counting mechanism is employed to count the number of codons (groups of three nucleotides) between this start codon and the next stop codon downstream of it. If a threshold value of 33 is reached, or exceeded, then the sequence encompassed by these two codons (start and stop codons) is defined as an ORF. This procedure is repeated, each time starting at the next nucleotide following the previous stop codon found, in order to identify all the other putative ORFs. The scan is performed on all three reading frames of both DNA strands of the phage sequence.

Sequence homology (BLAST) searches for each ORF are then carried out using an implementation of BLAST programs, although any of a variety of different sequence comparison and matching programs can be utilized as known to those skilled in the art. Downloaded public databases used for sequence analysis include:

i) non-redundant GenBank (ftp://ncbi.nlm.nih.gov/blast/db/nr.Z),

ii) Swissprot (ftp://ncbi.nlm.nih.gov/blast/db/swissprot.Z);

iii) vector (ftp://ncbi.nlm.nih.gov/blast/db/vector.Z);

iv) pdbaa databases (ftp://ncbi.nlm.nih.gov/blast/db/pdbaa.Z);

v) *Staphylococcus aureus* NCTC 8325 (ftp://ftp.genome.ou.edu/pub/staph/staph-1k.fa);

vi) streptococcus pyogenes (ftp://ftp.genome.ou.edu/pub/strep/strep-1k.fa);

vii) *streptococcus pneumoniae* (ftp://ftp.tigr.org/pub/data/s_pneumoniae/gsp.contigs.112197.Z);

viii) *mycobacterium tuberculosis* CSU#9 (ftp://ftp.tigr.org/pub/data/m_tuberculosis/TB_091097.Z) and ix) *pseudomonas aeruginosa* (http://www.genome.washington.edu/pseudo/data.html).

The results of the homology searches performed on the ORFs is shown in Table 4.

Example IV. Subcloning of Bacteriophage 44AHJD ORFs into a Staph A inducible expression system.

Preparation of shuttle vectors

The shuttle vector pT0021, in which the firefly luciferase (lucFF) expression is controlled by the ars (arsenite) promoter/operator (Tauriainen *et al.*, 1997), was modified in the following fashion. Two oligonucleotides corresponding to a short antigenic peptide derived from the heamaglutinin protein of influenza virus (HA epitope tag) were synthesized (Field *et al.*, 1988). The sense strand HA tag sequence (with BamHI, SalI and HindIII cloning sites) is:

5'-gatcccggtcgaccaagcttTACCCATACGACGTCCCAGACTACGCCAGCTGA-3'

(where upper case letters denote the nucletotide sequence of the HA tag); the antisense strand HA tag sequence (with a HindIII cloning site) is:

5'-agctTCAGCTGGCGTAGTCTGGGACGTCGTATGGGTAaagcttggtcgaccgg-3'

(where upper case letters denote the sequence of the HA tag). The two HA tag oligonucleotides were annealed and ligated into pT0021 vector which had been digested with BamHI and HindIII. This manipulation resulted in replacement of the lucFF gene by the HA tag. This modified shuttle vector containing the arsenite inducible promoter, the arsR gene, and HA tag was named pTHA. A diagram outlining our modification of pT0021 to generate pTHA is shown in Fig. 1A.

The shuttle vector pT0021 was also modified as below to suit our specific application. Two oligonucleotides were synthesized. The sense strand sequence (with XhoI cloning site) is: 5'-AATTCTCGAGTAAAATAACAT-3' ; the antisense strand sequence (with a BamHI cloning site) is:

5'-CGGGATCCGCCTCCTTTTCTCAACAGTCACCTGATTT -3'. The two oligonucleotides were used for polymerase chain reaction (PCR) amplification of pT0021 vector. The PCR product was gel purified using the Qiagen kit as described, and digested with XhoI and BamHI. The digested PCR product was again gel purified, ligated into XhoI and BamHI digested pT0021 vector, and used to transform *E. coli* bacterial strain DH10β (as described above). This manipulation results in the construction of a pT0021-intermediated vector containing a RBS sequence located immediately upstream of the BamHI cloning site. Two other oligonucleotides were synthesized. The sense strand sequence (with BamHI cloning site) is:

5'-CGGGATCCATGAGGGGTTCCGAAGACG-3' ; the antisense strand sequence (with a HindIII cloning site) is: 5'-CCCAAGCTTACAATTTGGACTTTC -3'. The two oligonucleotides were used for PCR amplification of pT0021-intermediated vector. The PCR product was gel purified and digested with BamHI and HindIII. The digested PCR product was then gel purified as described, ligated into BamHI and HindIII digested pT0021-intermediated vector, and used to transform *E. coli* bacterial strain DH10β. This modified shuttle vector containing the ATG of the lucFF gene located immediately downstream of the BamHI cloning site was named pTM. A diagram outlining our modification of pT0021 to generate pTM is shown in Fig 1B.

As another example of inducible promotor, the arsenite-inducible promotor and the asrR gene from the pTM vector were replaced by a lactose-inducible promotor and the lacR gene from *Staphylococcus aureus*. The *S. aureus* gene encoding for the repressor of the lac operon (lacR) is found immediately upstream of the promoter-proximal end of the the lacA-G genes. Two oligonucleotides corresponding to a 2.18kb-DNA region encompassing the lacR and the lac operon promotor region were synthesized. The sense strand sequence is : 5'-ccgctcgagCTCCAAATTCCAAAACAG-3' (with a XhoI cloning site, ctcgag) ; the antisense strand sequence is: 5'-cgggatccAATAAGACTCCTTTTTAC-3'(with a BamHI

cloning site, ggatcc). These two oligonucleotides were used for the PCR amplification of *Staphylococcus aureus* DNA. The PCR product was gel purified and digested with XhoI and BamHI. The digested PCR product was also gel purified, ligated into XhoI and BamHI-digested pTM vector, and used to transform *E. coli* bacterial strain DH10β. In the resulting vector, pTMLac, the firefly luciferase (lucFF) expression is under the control of the *S. aureus* lac operon promoter/operator. Recombinant pTMLac clones were picked and the sequence integrity of the 2.18kb-lac operon region (lacR + lac promotor) was verified directly by DNA sequencing. A diagram outlining the pTMSLac vector characteristics is shown in Fig 1C.

Cloning of ORFs with a Shine-Dalgarno sequence.

Each ORF, encoded by Bacteriophage 44AHJD, larger than 33 amino acids and having a Shine-Dalgarno sequence upstream of the initiation codon was selected for functional analysis for bacterial inhibition. In total, 31 ORFs were selected and screened as detailed below. A list of these is presented in Fig 4A. As outlined in Fig 2A, each individual ORF, from initiation codon to last codon (excluding the stop codon), was amplified from phage genomic DNA using the polymerase chain reaction (PCR). For PCR amplification of ORFs, each sense strand primer targets the initiation codon and is preceded by a BamHI restriction site (5'cgggatcc3') and each antisense oligonucleotide targets the pentultimate codon (the one before the stop codon) of the ORF and is preceded by a Sal I restriction site (5'gcgtcgaccg3'). The PCR product of each ORF was gel purified and digested with BamHI and SalI. The digested PCR product was then gel purified using the Qiagen kit as described, ligated into BamHI and SalI digested pTHA vector, and used to transform *E. coli* bacterial strain DH10β (as described above). As a result of this manipulation, the HA tag is set inframe with the ORF and is positioned at the carboxy terminus of each ORF (pTHA/ORF clones). Recombinant pTHA/ORF clones were picked and their insert sizes were confirmed by PCR analysis using primers flanking the cloning site. The names and sequences of the primers that were used for the PCR amplification were: HAF: 5'TATTATCCAAAACTTGAACA3'; HAR: 5'CGGTGGTATATCCAGTGATT3'. The sequence integrity of cloned ORFs was verified directly by DNA sequencing using primers

HAF and HAR. In cases where verification of ORF sequence could not be achieved by one pass with the sequencing primers, additional internal primers were selected and used for sequencing.

Each ORFs cloned into pTHA were also tested following removal of the HA tag. The pT/ORF vectors were obtained by Hind III digestion of individual pTHA/ORFs, gel purification of vector and religation of Hind III ends together.

ORF 12 and ORF 25 were also clones into pTM and pTMLac respectively. Each individual ORF, from initiation codon to stop codon was amplified from phage genomic DNA using the PCR. Each sense strand primer targets the initiation codon and is preceded by a BamHI restriction site (5'-cgggatcc-3') and each antisense oligonucleotide targets the stop codon of the ORF and is preceded by a HindIII restriction site (5'-cccaagctt-3'). The PCR product of each ORF was purified using the Quiagen kit as described and digested with BamHI and HindIII. The digested PCR product was also purified using the Quiagen kit, ligated into BamHI and HindIII digested pTM or pTMLac vector and used to transform *E. coli* bacterial strain DH10β (as described above). As a result of this manipulation, the ORF is under the control of the arsenite-inducible (pTM) or lactose-inducible (pTMLac) promotors. Recombinant clones were picked and their insert sizes were confirmed by PCR analysis using primers flanking the cloning site. The sequence integrity of cloned ORFs was verified directly by DNA sequencing using primers HAF and HAR.

Example V. Functional assay for bacterial inhibitory activity of bacteriophage 44AHJD ORFs.

Transformation of *Staphylococcus aureus* with expression construct

Staphylococcus aureus strain RN4220 (Kreiswirth *et al.*, 1983) was used as a recipient for the expression of recombinant plasmids. Electoporation was performed essentially as previously described (Schenk and Laddaga, 1992). Selection of recombinant clones was performed on Luria-Broth agar (LB-agar) plates containing 30 μg/ml of kanamycin.

For each ORF introduced in the pTHA and pT plasmids, 3 independent transformants were isolated and used to individually inoculate cultures in 5 ml of TSB containing 30μg/ml kanamycin, followed by growth to saturation (16 hrs at 37°C). An aliquot of this stationary phase culture was used to generate a frozen glycerol stock of the transformant (stored at - 80°C)

The presence of individual phage 44AHJD ORF DNA inserts in the plasmid was verified by PCR amplification using 1.5 μl transformant miniprep DNA in a PCR with primers flanking the cloning site of ORF in pTHA vector (HAF and HAR). The composition of the PCR reaction and the cycling parameters are identical to those employed for library screening described above.

Induction of gene expression from the ars- and lac-inducible promotors

Sodium arsenite ($NaAsO_2$) was purchased from Sigma (Sigma-Aldrich Canada LTD, Oakville) and was used as heavy metals to induce gene expression from the ars promoter/operator in solid and liquid medium assays.

The lactose (lac) genes of *Staphylococcus aureus* have been shown to be inducible with the addition of either lactose or galactose to the culture medium (Oskouian & Stewart, 1990, J. Bacteriol. 172 : 3804-3812). Galactose (2%w/v) was used to induce the gene expression from the lac promotor/operator in liquid assay.

At pre-determined times, appropriated inducer was added to the culture to induce transcription of the phage ORFs cloned immediately downstream from an arsenite-inducible promoter in the expression plasmids pTHA, pT, or pTM, or a lactose-inducible promotor in the expression plasmid pTMLac. The anti-microbial activity of individual phage 44AHJD ORFs was monitored by two growth inhibitory assays, one on solid agar medium, the other in liquid medium.

The effect of ORF induction on bacterial growth characteristics was then monitored and quantitated.

a Screening on semi-solid support media

ORFs cloned into pTHA and pT vectors were first screened by the functional assay on semi-solid medium as outlined in Fig 3A. Cells containing different recombinant

plasmids were grown overnight at 37°C in LB medium supplemented with 30 μg/ml of kanamycin. The cells were then diluted and the identification of inhibitory ORFs was performed by spotting 3 ul of each dilution of *S. aureus* transformed cells containing phage 44AHJD ORFs onto agar plates containing increasing concentrations of sodium arsenite (0; 2.5; 5; and 7.5 μM) and Kanamycin. The plates were incubated overnight at 37°C, after which a growth inhibition of the ORF transformants on plates that contain arsenite are compared to plates without arsenite. Noninduced and induced cultures of S *aureus* transformed with a non-inhibitory ORF (77 bacteriophage ORF 30 cloned into pT vector) were included as negative control. The 77 ORF 30 amino acids residue composition from N-terminal to C-terminal is:

MKIKVKKEMRLDELIKWARENPDLSQGKIFFSTGFSDGFVRFHPNTNKCSTS SFIPIDIPFIVDIEKEVTEETKVDRLIELFEIQEGDYNSTLYENTSIKECLYGRCVPTKAF YILNDDL TMTLIWKDGELLV.

Results of the 31 bacteriophage ORFs tested for functional assay on semi-solid media are listed in Fig 4A. Among them, induction of expression of phage 44AHJD OERF12 and 25 results in the inhibition of growth of the *S. aureus* transformants. Fig 4B shows the result of growth inhibition with three clones of *S. aureus* expressing these inhibitory ORFs or the control non-inhibitory 77 ORF 30.

b Quantification of growth inhibition in liquid medium

As outlined in Fig 3B, the effect of ORF induction on bacterial growth inhibition was then further quantitated by functional assay in liquid medium. Cells containing phage 44AHJD ORF 12 or 25 were grown for overnight at 37°C in LB medium supplemented with the appropriate antibiotic selection. These cultures were 50-fold dilution with fresh media containing kanamycin and the growth was continued for 2 h at 37°C. The same OD565 equivalent of cultures (approximately 1 ml) was added to 19 ml of fresh media containing kanamycin and transferred to a 125 ml-Erlenmeyer flask. The cultures were incubated for an additional 4 hrs at 37°C in the absence or in the presence of inducer (sodium arsenite at the final concentrations of 5.0 μM or 2.0% galactose). During that period of time, the effect of expression of the phage 44AHJD ORFs on bacterial cell growth was monitored, at each time

point intervals, by measuring the OD_{565} and the number of colony forming units (CFU) in the cultures containing or not the inducer. The number of CFU was evaluated as followed. Cultures were serially diluted and aliquots from induced and uninduced cultures were plated out on agar plates containing an appropriate antibiotic selection but lacking inducer. Following incubation overnight at 37°C, the number of colonies was counted. Cultures of S *aureus* transformed with a non-inhibitory ORF (77 bacteriophage ORF 30 cloned into pT vector or 44AHJD ORF 114 cloned into pTM) were included as control.

As shown in Fig 5, for each inhibitory ORFs, the number of CFU and OD increased over time under non-induced conditions. Similar growth rates were also observed with transformants harboring non-inhibitory ORF under both induced and non-induced conditions. Cultures of *S. aureus* transformants harboring the phage ORF 12 or 25 shown a significant lower growth rate compared to their respective parallel cultures grown under noninduced conditions. Induction of expression of ORF 12 or 25 were cytocydal for the bacterial growth. As shown in Fig 5A, the expression of ORF 12 results in a rapid decrease in the number of CFU. A one log reduction in the number of CFU compared to the number of CFU initially present in the same culture was observed at 1h following induction with sodium arsenite.

As shown in Fig 5B, the expression of ORF 25 results in a 2 log reduction in the number of CFU compared to the number of CFU initially present in the same culture. The induction of the expression of the same ORF with galactose results in a half log reduction in the number of CFU compared to the number of CFU initially present in the same culture was observed.

Example VI. Phage ORF protein expression analysis in *S. aureus*

The level of expression of the inhibitory ORFs was measured by performing Western blot analyses. *Staphylococcus aureus* strain RN4220 was electroporated with each inhibitory ORFs cloned into pTHA vector as described above. Cells containing different recombinant plasmids were grown for overnight at 37°C in TSB (Tryptic soy broth, DIFCO) medium in the presence of 30 μg/ml kanamycin. The overnight cultures were subjected to a 50-fold dilution with fresh media containing kanamycin and the growth was continued for 2 h at

37°C. At the end, cells were diluted with fresh TSB medium containing or not 5.0 μM of Sodium Arsenite, in the presence of kanamycin and incubated at 37°C for an additional 3.5 h. The same OD_{565} equivalent of cultures was centrifuged at 3000 g for 5 min and washed with 20 ml of TBS buffer (140 mM NaCl, 25 mM Tris-HCl, pH 7.5) containing protease inhibitors (1 mM of each phenylmethylsulfonyl fluoride (PMSF) and N-ethylmalemyde (NEM)). For lysis, cell pellets were resuspend in 25 μl with TBS buffer containing 1 mM PMSF, 1 mM NEM, 20 μg of each DNAse I and RNase A and 50 Units/ ml of lysostaphin, and incubated at 37°C for 1 h. The reaction was stopped by the addition of 25 μl of 2x SDS buffer (100 mM Tris pH 6.8, 4% SDS, 200 mM DTT, 20% Glycerol and 0.2% Bromophenol blue). Cell lysates were boiled for 10 min, centrifuged for 10 min at 13,000 g and 10-15 μl of the lysates were loaded onto a 15-18% SDS-page using Tris-Glycine-SDS as a running buffer (3.03 g of Tris HCl, 14.4 g of Glycine and 0.1 % SDS per liter). After migration, proteins were transferred onto an immobilon-P membrane (PVDF, Millipore) using Tris-Glycin-Methanol as a transfer buffer (3.03 g Tris, 14.4 Glycine and 200 ml Methanol per liter) for 2 hrs at 4°C at 100 V. PVDF membrane was pretreated in methanol for 30 s, washed 4-5 times with H_2O and soaked in transfer buffer.

After the transfer, the membrane was blocked in 20 ml of TBS containing 0.05 % Tween-20 (TBST), 5% skim milk and 0.5% gelatin for 1 hr at room temperature and then, a pre-blocking antibody (ChromPureRabbit IgG, Jackson immunoResearch lab. # 011-000-003) was added at a dilution of 1/750 and incubated for 1 hr at room temperature or O/N at 4°C. Membrane was washed 6 times for 5 min in TBST at room temperature. The primary antibody (murine mono-HA antibody, Babco # MMS-101 P) directed against the HA epitope tag and diluted 1/1000 was then added and incubated for 3 h at room temperature in the presence of 5% Skim Milk and 0.5% Gelatin. Membrane was washed 6 times for 5 min in TBST at room temperature. A secondary antibody (anti-mouse IgG, peroxidase-linked species-specific whole antibody, Amersham # NA 931) diluted 1/1500 (7.5 μl in 10 ml) was then added and incubated for 1 hr at room temperature. After 6 washes in TBST, the membrane was briefly dried and then, the substrate (Chemiluminescence reagent plus, Mandel # NEL104) was added to the membrane and incubated for 1 min at room temperature. The membrane was briefly dried and exposed to x-ray film (Kodak, Biomax

MS/MR) for different periods of time (30 s to 10 min). As shows in Figure 6, the presence of sodium arsenite in the cultures induces the expression of proteins corresponding to the phage 44AHJD ORF 12 and 25.

Example VII. Screening Assays

Phage display

Phage display is a powerful assay to measure protein:protein interaction. In this scheme, proteins or peptides are expressed as fusions with coat proteins or tail proteins of filamentous bacteriophage. A comprehensive monograph on this subject is Phage Display of Peptides and Proteins. A Laboratory Manual edited by Kay et al. (1996) Academic Press. For phages in the Ff family that include M13 and fd, gene III protein and gene VIII protein are the most commonly-used partners for fusion with foreign protein or peptides. Phagemids are vectors containing origins of replication both for plasmids and for bacteriophage. Phagemids encoding fusions to the gene III or gene VIII can be rescued from their bacterial hosts with helper phage, resulting in the display of the foreign sequences on the coat or at the tip of the recombinant phage.

In the simplest assay, purified recombinant dnaN protein, or a fragment of dnaN, could be immobilized in the wells of a microtitre plate and incubated with phages displaying 44AHJD ORF 25 in fusion with the gene III protein. Washing steps are performed to remove unbound phages and bound phages are detected with monoclonal antibodies directed against phage coat protein (gene VIII protein). Color development by means of an enzyme-linked secondary antibody allows quantitative detection of bound fusion protein. Screening for inhibitors is performed by the incubation of the compound with the immobilized target before the addition of phages. The presence of an inhibitor will specifically reduce the signal in a dose-dependent manner relative to controls without inhibitor.

Identification of the surface of interaction on both polypeptide partners.

The invention provides a method for the identification of 44AHJD ORF 25 and DnaN polypeptide fragments which are involved in the interaction between these two proteins. These fragments may include, for example, truncation polypeptides having a portion of an amino acid sequence of any of the two proteins, or variants thereof, such as a

continuous series of residues that includes an amino- and/or carboxyl-terminal amino acid sequence.

Partial proteolysis of proteins in solution is one method to delineate the domain boundaries in multi-domain proteins. By subjecting proteins to limited digestion, the most accessible cleavage sites are preferentially hydrolyzed. These cleavage sites preferentially reside in less structured regions which include loops and highly mobile areas typical of the joining amino acids between highly structures domains. Purified 44AHJD ORF 25 or DnaN proteins can be subjected to partial proteolysis. The proteolysis can be performed with low concentrations of proteases (trypsin, chymotrypsin, endoproteinase Glu-C, and Asp-N) with 44AHJD ORF 25 or DnaN in solution, resulting in the generation of defined proteolytic products as observed by SDS-PAGE. An acceptable concentration and reaction time is defined by the near complete conversion of the full-length protein to stable proteolytic products. The proteolytic products are then subjected to affinity chromatography containing the appropriated partner of interaction (44AHJD ORF 25 or DnaN purified proteins) to determine a protein sub-region able to interact. Interacting domains are identified by mass spectrometry to determine both the intact fragment mass and the completely digested with trypsin (by in-gel digestion) to better determine the amino acid residues contained within the partial proteolytic fragment. Using both sets of data, the amino acid sequence of the partial proteolytic fragment can be precisely determined.

Another approach is based on peptide screening using different portions of 44AHJD ORF 25 and DnaN to identify minimal peptides from each polypeptide that are able to disrupt the interaction between the two proteins. It is assumed that fragments able to prevent interaction between 44AHJD ORF 25 and DnaN correspond to domains of interaction located on either of the two proteins. The different peptide fragments can be screened as competitors of interaction in protein: protein binding assays such as the ones described above. Fine mapping of interaction site(s) within a protein can be performed by an extensive screen of small overlapping fragments or peptides spanning the entire amino acid sequence of the protein.

Fragments of 44AHJD ORF 25 or of DnaN can be produced by proteolytic digestion of the full-length proteins as descibed above. Alternatively, suitable dnaN or 44AHJD ORF

25-derived amino acid fragments representative of the complete sequence of both proteins can be chemical synthesis. For instance, in the multipin approach, peptides are simultaneously synthesis by the assembly of small quantities of peptides (ca. 50 nmol) on plastic pins derivatized with an ester linker based on glycolate and 4-(hydroxymethyl) benzoate (Maeji 1991 Pept Res, 4:142-6).

Functional assays for bacterial growth: OD and CFU measurement over time

Compounds selected for their ability to inhibit the 44AHJD ORF 25-DnaN interaction can be further tested in functional assays on bacterial growth. Cells are grown in the presence of varying concentrations of a candidate compound added directly to the medium. The cultures are then incubated for an additional 4 hrs at 37°C. During that period of time, the effect of inhibitors on bacterial cell growth may be monitored, at 40 min intervals, by measuring the OD565 and the number of colony forming units (CFU) in the cultures. The number of CFU is evaluated as follows: cultures are serially diluted and aliquots from the different cultures are plated out on agar plates. Following incubation overnight at 37°C, the number of colonies are counted. Non-treated cultures of *S. aureus* are included as control.

Surface plasmon resonance

Another powerful assay to screen for inhibitors of a protein: protein interaction is surface plasmon resonance. Surface plasmon resonance is a quantitative method that measures binding between two (or more) molecules by the change in mass near the sensor surface caused by the binding of one protein or other biomolecule from the aqueous phase to a second protein or biomolecule immobilized on the sensor. This change in mass is measured as resonance units versus time after injection or removal of the second protein or biomolecule and is measured using a Biacore Biosensor (Biacore AB). dnaN could be immobilized on a sensor chip (for example, research grade CM5 chip; Biacore AB) using a covalent linkage method (e.g. amine coupling in 10 mM sodium acetate [pH 4.5]). A blank surface is prepared by activating and inactivating a sensor chip without protein immobilization. The binding of 44AHJD ORF 25 to dnaN, or a fragment of dnaN, is measured by injecting purified 44AHJD ORF 25 over the chip surface. Measurements are

performed at room temperature. Conditions used for the assay (i.e., those permitting binding) are as follows: 25 mM HEPES-KOH (pH 7.6), 150 mM sodium chloride, 15% glycerol, 1 mM dithiothreitol, and 0.001% Tween 20 with a flow rate of 10 ul/min. Preincubation of the sensor chip with candidate inhibitors will predictably decrease the interaction between 44AHJD ORF 25 and dnaN. A decrease in 44AHJD ORF 25 binding is indicative of competitive binding by the candidate compound.

Fluorescence Resonance Energy Transfer (FRET)

Another method of measuring inhibition of binding of two proteins uses fluorescence resonance energy transfer (FRET; de Angelis, 1999, Physiological Genomics). FRET is a quantum mechanical phenomenon that occurs between a fluorescence donor (D) and a fluorescence acceptor (A) in close proximity (usually < 100 A of separation.) if the emission spectrum of D overlaps with the excitation spectrum of A. Variants of the green fluorescent protein (GFP) from the jellyfish Aequorea victoria are fused to a polypeptide or protein and serve as D-A pairs in a FRET scheme to measure protein-protein interaction. Cyan (CFP: D) and yellow (YFP: A) fluorescence proteins are linked with dnaN polypeptide, or a fragment of dnaN and 44AHJD ORF 25 protein respectively. Under optimal proximity, interaction between dnaN, or a fragment of dnaN, and 44AHJD ORF 25 causes a decrease in intensity of CFP concomitant with an increase in YFP fluorescence.

The addition of a candidate modulator to the mixture of appropriately labeled dnaN and 44AHJD ORF 25 protein, will result in an inhibition of energy transfer evidenced by, for example, a decease in YFP fluorescence at a given concentration of 44AHJD ORF 25 relative to a sample without the candidate inhibitor.

Fluorescence polarization

In addition to the surface plasmon resonance and FRET methods, fluorescence polarization measurement is useful to quantitate protein-protein binding. The fluorescence polarization value for a fluorescently-tagged molecule depends on the rotational correlation time or tumbling rate. Protein complexes, such as those formed by *S. aureus* dnaN polypeptide, or a fragment of dnaN associating with a fluorescently labeled polypeptide (e.g., 44AHJD ORF 25 or a binding fragment thereof), have higher polarization values than a fluorescently labeled monomeric protein. Inclusion of a candidate inhibitor of the dnaN

interaction results in a decrease in fluorescence polarization relative to a mixture without the candidate inhibitor if the candidate inhibitor disrupts or inhibits the interaction of dnaN with its polypeptide binding partner. It is preferred that this method be used to characterize small molecules that disrupt the formation of polypeptide or protein complexes.

Bio Sensor Assay

ICS biosensors have been described by AMBRI (Australian Membrane Biotechnology Research Institute; http//www.ambri.com.au/). In this technology, the self-association of macromolecules such as dnaN, or a fragment of dnaN, and bacteriophage 44AHJD ORF 25, is coupled to the closing of gramacidin-facilitated ion channels in suspended membrane bilayers and hence to a measurable change in the admittance (similar to impedence) of the biosensor. This approach is linear over six order of magnitude of admittance change and is ideally suited for large scale, high through-put screening of small molecule combinatorial libraries.

Example VIII: Identification of bacterial target

To identify the *S. aureus* protein(s) that interact with inhibitory ORF 25 of *S. aureus* bacteriophage 44AHDK, a GST-fusion of 44AHJD ORF 25 was generated. The recombinant protein was purified and utilized to make a GST/44AHJD ORF 25 affinity column. Cellular extracts prepared from *S. aureus* cells were incubated with the affinity matrix and the matrix was washed with buffers containing increasing concentrations of salt and different detergents. The protein elution profile was assessed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). A protein of molecular mass ~ 48 kDa, identified as PT48, was spicifically eluted from the affinity matrix and was not detected in eluates from GST negative control column. Eluted proteins were further characterized to determine the identity of the interacting protein.

Generation of GST/ORF 25 recombinant protein

Bacteriophage 44AHJD ORF 25 was sub-cloned into pGEX 4T-1 (Pharmacia), an expression vector containing the GST moiety. ORF 25 was obtained by digestion of pTHA/ 44AHJD ORF 25 (Figure 2A) with *BamHI* and *SalI*. The DNA fragment containing ORF

25 was gel purified by QiaQuick spin columns (Qiagen) and ligated into pGEX 4T-1 (which had been previously digested with Bam HI and Sal I) to generate pGEX 4T/ORF 25. Recombinant expression vectors were identified by restriction enzyme analysis of plasmid minipreps. Large-scale DNA preparations were performed and the resulting insert was sequenced. Test expressions in *E. coli* BL21 (DE3) Gold cells containing the expression plasmids were performed to identify optimal protein expression conditions. *E. coli* cells containing the expression constructs were grown in Luria-Bertani Broth at 25°C to an OD_{600} of 0.4 to 0.6 and induced with 1 mM IPTG for the optimal times and at the optimal temperatures (typically a 2 liter culture of BL21 (DE3) Gold (pGEX 4T/ORF25) grown at 25°C for 3hrs).

Fusion protein purification.

Cells containing GST/ORF 25 fusion protein were suspended in 15 ml lysis buffer/liter of cell culture with GST lysis buffer (20 mM Hepes pH 7.2, 500 mM NaCl, 10 % glycerol, 1 mM DTT, 1mM EDTA, 1mM benzamidine, and 1 PMSF) and lysed using a French pressure cell followed by three bursts of twenty seconds with an ultra-sonicator at 4°C. Triton X-100 was added to the lysate to a final concentration of 0.1% and mixed for 30 minutes at 4°C. The lysate was centrifuged at 4°C for 30 minutes at 10,000 rpm in a Sorval SS34 rotor. The supernatant was applied to a 4ml glutathione sepharose column pre-equilibrated with lysis buffer and allowed to flow by gravity. The column was washed with 10 column volumes of lysis buffer and eluted in 1.5 ml fractions with GST elution buffer (20 mM Hepes pH 8.0, 500 mM NaCl, 10 % glycerol, 1 mM DTT, 0.1mM EDTA, and 25 mM reduced glutathione). The fractions were analyzed by SDS-12.5% PAGE (Laemmli) and proteins were visualized by staining with Coomassie Brilliant Blue R250 stain to assess the amount of eluted GST/ORF 25 protein.

Affinity column preparation.

GST and GST/ORF25 were dialyzed overnight against affinity chromatography buffer (ACB; 20 mM Hepes pH 7.5, 10 % glycerol, 1 mM DTT, and 1 mM EDTA) containing 1 M NaCl. Protein concentrations were determined by Bio-Rad Protein Assay

and crosslinked to Affigel 10 resin (Bio-Rad) at protein/resin concentrations of 0, 0.1, 0.5, 1.0, and 2.0 mg/ml. The crosslinked resin was sequentially incubated in the presence of ethanolamine, and bovine serum albumin (BSA) prior to column packing and equilibration with ACB containing 100 mM NaCl.

S. aureus extract preparation.

Two extracts were prepared from *S. aureus* cell pellets. One lysate was prepared by French pressure cell lysis followed by sonication, and the other by lysostaphin-mediated digestion followed by sonication. The French pressure cell lysate was prepared by suspending 3 g of frozen *S. aureus* cells in ABC containing 500 mM NaCl, 1 mM PMSF, and 1 mM benzamidine. The suspended cells were subjected to three passes through the French pressure cell followed by 3 sonication bursts of 20 seconds each, made up to 0.1% Triton X-100, stirred for 30 minutes, and centrifuged at 50,000 rpm for 3 hrs in a Ti70 fixed angle Beckman rotor. The efficiency of cell lysis was low and the resulting lysate (7 ml) contained 2.4 mg/ml protein. The pellet after French pressure cell lysis was subjected to cryogenic grinding in liquid nitrogen in the same buffer with a mortar and pestle. The lysate was made up to 0.1% Triton X-100, stirred for 30 minutes, and centrifuged at 50,000 rpm for 3 hrs in a Ti70 fixed angle Beckman rotor yielding a lysate (10 ml) containing 2.0 mg/ml protein. The cell lysates were pooled, concentrated to 8 ml, and dialyzed overnight in a 3000 Mr cut-off dialysis membrane against ACB containing 1 mM PMSF, 1 mM benzamidine, and 75 mM NaCl. The dialyzed protein extract was removed from the dialysis tubing, centrifuged at 10 000 rpm in a Sorval SS34 rotor for 1 hr, and assayed for protein content (Bio-Rad Protein Assay) and salt concentration (conductivity meter).

Affinity chromatography.

The *S. aureus* extract was centrifuged at 4°C in a micro-centrifuge for 15 minutes and 200µl was applied to 20µl columns containing 0, 0.1, 0.5, 1.0, and 2.0 mg/ml ligand. ACB containing 100 mM NaCl (200µl) was applied to a control column containing 2.0 mg/ml ligand. The columns were washed with 10 column volumes ACB containing 100 mM NaCl and sequentially eluted with ACB containing 1% Triton X-100 and 100 mM NaCl

(800µl), ACB containing 1 M NaCl (800µl), and 1% SDS (160µl). 40µl of each eluate was resolved by SDS-12.5% PAGE (Laemmli) and the protein was visualized by silver stain.

Identification of *S. aureus* DnaN as an 44AHJD ORF 25 interacting protein

Affinity chromatography was performed using GST and GST ORF25 as ligands coupled to Affigel 10 at protein/resin concentrations of 0, 0.1, 0.5, 1.0, and 2.0 mg/ml. Two *S. aureus* extracts were used for affinity chromatography with each of the ligands. Two extracts used for affinity chromatography, prepared separately, contained 4.0 and 9.0 mg/ml protein. One candidate interacting protein of 48 kDa (PT48) was observed in the 1% SDS eluates in the initial chromatography experiment (Figure 7B).

The candidate protein, PT48 was excised from the SDS-PAGE gels and prepared for tryptic peptide mass determination by MALDI-ToF mass spectrometry (Qin, J., Fenyo, D., Zhao, Y., Hall, W.W., Chao, D.M., Wilson, C.J., Young, R.A. and Chait, B.T. (1997) Anal. Chem. 69, 3995-4001). High quality mass spectra were obtained (Figure 8). The PT48 proteins observed in two affinity chromatography experiments were identical as determined by the masses of the tryptic peptides. Computational analysis (http://prowl.rockfeller.edu/cgi-bin/ProFound) of the mass spectrum obtained identifies the corresponding ORF in the *S. aureus* nucleotide sequence in the University of Oklahoma *S. aureus* genomic database (http://www.genome.ou.edu/staph.html). The identity of that protein which binds specifically to GST ORF25 is the DNA-directed DNA polymerase III beta subunit (Genbank accession #1084187) (Figure 9 and Figure 10).

It is important to note that in assays of protein-protein interaction, it is possible that a modulator of the interaction need not necessarily interact directly with the domain(s) of the proteins that physically interact. It is also possible that a modulator will interact at a location removed from the site of protein-protein interaction and cause, for example, a conformational change in the dnaN polypeptide. Modulators (inhibitors or agonists) that act in this manner are of interest since the change they induce may modify the activity of the dnaN polypeptide.

Compounds selected for their ability to bind to dnaN or to inhibit the 44AHJD ORF 25-dnaN interaction can be further tested in functional assays of bacterial growth. Cultures of *S. aureus* are grown in the presence of varying concentrations of a candidate compound added directly to the medium. The cultures are then incubated for an additional 4 hrs at 37°C. During that period of time, the effect of inhibitors on bacterial cell growth may be monitored at 40 min intervals, by measuring the OD565 and the number of colony forming units (CFU) in the cultures. The number of CFU is evaluated as follows: cultures are serially diluted and aliquots from the different cultures are plated out on agar plates. Following incubation overnight at 37°C, the number of colonies are counted. Non-treated cultures of *S. aureus* are included as negative control.

- REFERENCES

- Cohen, M.L. (1992). Science 257: 1050-1055.
- Rusterholtz, K., and Pohlschroder, M. (1999). Cell 96, 469-470.
- Ackermann, H.-W. and DuBow, M.S. (1987). Viruses of Prokaryotes. CRC Press. Volumes 1 and 2.
- Durfee, T., Becherer, K., Chen, P.-L., Yeh, S.-H., Yang, Y., Kilburn, A.E., Lee, W.-H., and Elledge, S.J. (1993). Genes & Dev. 7: 555-569.
- Sopta, M., Carthew, R.W., and Greenblatt, J. (1995) J. Biol. Chem. 260: 10353-10369.
- Qin, J., Fenyo, D., Zhao, Y., Hall, W.W., Chao, D.M., Wilson, C.J., Young, R.A. and Chait, B.T. (1997). Anal. Chem. 69: 3995-4001.
- Sambrook, J., Fritsch, E.F. and Maniatis, T (1989). Molecular cloning: A laboratory Manual. Cold Spring Harbor Laboratory, New York. Cold Spring Harbor Laboratory Press.
- Swanström, M. and Adams, M.H. (1951). Agar layer method for production of high titer phage stocks. Proc. Soc. Exptl. Biol. & Med. 78: 372-375.
- Tauriainen, S., Karp, M., Chang, W and Virta, M. (1997). Recombinant luminescent bacteria for measuring bioavailable arsenite and antimonite. Appl. Environ. Microbiol. 63:4456-4461.

- Field, J., Nikawa, J.-I., Broek, D., MacDonald, B., Rodgers, L., Wilson, I.A., Lemer, R.A., and Wigler, M. (1988). Purification of a RAS-responsive adenylyl cyclase complex from Saccharomyces cerevisiae by use of an epitope addition method. Mol. Cell. Biol. 8: 2159-2165.
- Kreiswirth, BN., Lofdahl, S., Belley, MJ., O'Reilly, M., Shlievert, PM., Bergdoll, MS. and Novicks, RP. 1983. Nature #305: 709-712.
- Schenk, S. and Laddaga, RA. 1992. FEMS Microbiology Letters #94: 133-138.
- Oskouian, B. and Stewart, GS. 1990. J. Bacteriol. #172: 3804-3812.

All patents and publications mentioned in the specification are indicative of the levels of skill of those skilled in the art to which the invention pertains. All references cited in this disclosure are incorporated by reference to the same extent as if each reference had been incorporated by reference in its entirety individually.

One skilled in the art would readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The specific methods and compositions described herein as presently representative of preferred embodiments are exemplary and are not intended as limitations on the scope of the invention. One of ordinary skill in the art would recognize that Bacteriophage 44AHJD ORFs described herein are provided and discussed by way of example, and other the ORFs of Bacteriophage 44AHJD, including amino acid sequences and nucleic acid sequences which encode products, are within the scope of the present invention. Changes therein and other uses will occur to those skilled in the art which are encompassed within the spirit of the invention are defined by the scope of the claims.

It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention. For example, those skilled in the art will recognize that the invention may suitably be practiced using a variety of different expression vectors and sequencing methods within the general descriptions provided.

The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations which is not specifically disclosed

herein. Thus, for example, in each instance herein any of the terms "comprising," "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is not intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims.

In addition, where features or aspects of the invention are described in terms of Markush groups or other grouping of alternatives, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group or other group. For example, if there are alternatives A, B, and C, all of the following possibilities are included: A separately, B separately, C separately, A and B, A and C, B and C, and A and B and C.

Thus, additional embodiments are within the scope of the invention and within the following claims.